

150517

**Shears, Beverly**

**From:** Devi, Sarvamangala  
**Sent:** Tuesday, April 12, 2005 9:21 AM  
**To:** Shears, Beverly  
**Subject:** 09/077,574

Good morning Beverly:

In application 09/077,574, please perform a sequence search for SEQ ID NO: 2 in both commercial and pending databases.

Thanx.

S. DEVI, Ph.D.  
 Primary Examiner  
 AU 1645  
 Rems - 3C18

1

Date completed: \_\_\_\_\_

Searcher: Beverly e 2528

Terminal time: \_\_\_\_\_

Elapsed time: \_\_\_\_\_

CPU time: \_\_\_\_\_

Total time: \_\_\_\_\_

Number of Searches: \_\_\_\_\_

Number of Databases: \_\_\_\_\_

**Search Site**

\_\_\_\_\_ STIC

\_\_\_\_\_ CM-1

\_\_\_\_\_ Pre-S

**Type of Search**

\_\_\_\_\_ N.A. Sequence

\_\_\_\_\_ A.A. Sequence

\_\_\_\_\_ Structure

\_\_\_\_\_ Bibliographic

**Vendors**

\_\_\_\_\_ IG

\_\_\_\_\_ STN

\_\_\_\_\_ Dialog

\_\_\_\_\_ APS

\_\_\_\_\_ Geninfo

\_\_\_\_\_ SDC

\_\_\_\_\_ DARC/Questel

\_\_\_\_\_ ☒ Other CGN

**THIS PAGE BLANK (USPTO)**



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 130517

**TO: Sarvamangala Devi**  
**Location: REM/3C18**  
**Wednesday, April 13, 2005**  
**Art Unit: 1645**  
**Serial Number: 09/077574**

**From: Beverly Shears**  
**Location: Biotech-Chem Library**  
**REM 1A54**  
**Phone: 571-272-2528**  
**beverly.shears@uspto.gov**

### Search Notes

#### Protein Sequence Searches – February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rup) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (uniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 09:29:17 ; Search time 70 Seconds  
(without alignments)  
4008.848 Million cell updates/sec

Title: US-09-077-574A-2

Perfect score: 2708

Sequence: 1 MASKEILFDKAREKLSRGV.....KDMPPGCGMGGMGDMGY 548

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2708	100.0	548	2 O87888	O87888 lawsonia in
2	2462.5	90.9	547	2 Q72AL6	Q72AL6 desulfovibr
3	1967.5	72.7	544	2 Q747C7	Q747C7 geobacter s
4	1900.5	70.2	547	2 O6N0B9	O6N0B9 magnetospir
5	1891.5	69.8	547	1 CH60_RALSO	O8Y1P8 talstonia s
6	1880	69.4	545	1 CH60_ACEAC	Q8YB22 acetobacter
7	1878.5	69.4	547	2 Q75T66	Q75T66 burkholderi
8	1874.5	69.2	547	2 Q6Z2C7	Q6Z2C7 myxococcus
9	1872.5	69.1	546	1 CH60_BURCE	Q2ZFE0 burkholderi
10	1869.5	69.0	547	1 CH61_RHOPA	P60364 rhodopseudo
11	1868.5	69.0	546	1 CH60_BURPS	O9F712 burkholderi
12	1868.5	69.0	546	1 CH60_BURVI	Q9ZFD8 burkholderi
13	1868.5	69.0	550	2 Q621B2	Q621B2 burkholderi
14	1868	69.0	546	1 CH60_RHOPA	Q93MH1 rhodopseudo
15	1867.5	69.0	546	1 CH60_BURTH	P58723 burkholderi
16	1867.5	69.0	546	2 Q63RH5	Q63RH5 burkholderi
17	1867.5	69.0	547	1 CH60_BARHE	O33963 bartonella
18	1867	68.9	545	1 CH61_RHIME	P35469 rhizobium m
19	1866	68.9	551	1 CH60_AMOPS	P26004 amceba prot
20	1864.5	68.9	547	1 CH60_BARQU	O33964 bartonella
21	1864	68.8	545	1 CH64_RHIME	Q92ZQ4 rhizobium m
22	1862.5	68.8	547	1 CH60_XYLFA	Q9FPF2 xyella fas
23	1862.5	68.8	547	1 CH60_XYLFT	Q87BC0 xyella fas
24	1854.5	68.5	545	2 Q7NQX1	Q7NQX1 chromobacte
25	1854	68.5	543	2 Q89DA6	Q89DA6 bradyrhizob
26	1853	68.4	544	2 Q75T68	Q75T68 burkholderi
27	1852.5	68.4	540	1 CH60_RHOMR	Q9XC9A rhodothermu
28	1852.5	68.4	546	1 CH60_CHRVI	P31293 chromatium
29	1851	68.4	539	2 Q83WK0	Q83WK0 burkholderi
30	1847	68.2	542	2 Q6W1C3	Q6W1C3 rhizobium s
31	1847	68.2	542	2 Q6W1D5	Q6W1D5 rhizobium s

32	1847	68.2	546	1 CH60_XANAC	Q8PP1 xanthomonas
33	1845	68.1	545	1 CH60_RHOCA	P35678 rhodobacter
34	1844.5	68.1	546	1 CH60_RHILV	P34939 rhizobium l
35	1844	68.1	546	1 CH60_XANCH	Q8IC7 xanthomonas
36	1843.5	68.1	547	1 CH60_BORBR	O7W84 bordetella
37	1843.5	68.1	547	1 CH60_BORPA	Q7W134 bordetella
38	1843.5	68.1	547	1 CH60_CORPE	P48210 bordetella
39	1843.5	68.1	552	1 CH60_COXBU	P19421 coxiella bu
40	1842.5	68.0	552	2 Q83Z44	Q83Z44 coxiella bu
41	1842	68.0	546	1 CH60_XANCP	Q8PD23 xanthomonas
42	1841	68.0	542	1 CH65_RHIME	P35471 rhizobium m
43	1840.5	68.0	544	1 CH60_AGR75	P30779 agrobacteri
44	1840	67.9	547	1 CH60_NITEU	Q82Y60 nitrosomona
45	1839.5	67.9	552	2 Q83Z43	Q83Z43 coxiella bu

ALIGNMENTS

RESULT 1  
O87888  
ID O87888 PRELIMINARY; PRT; 548 AA.  
AC O87888;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE GroEL/HSP60 homolog.  
OS Lawsonia intracellularis.  
OG Plasmid pISI-2.  
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;  
OC Desulfovibrionaceae; Lawsonia.  
OX NCBI\_TaxID=29546;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98386497; PubMed=9720028;  
RA Dale C.J.H., Moses E.K., Ong C.C., Morrow C.J., Reed M.B., Hasse D.,  
RA Strugnell R.A.;  
RT "Identification and sequencing of the groE operon and flanking genes  
of Lawsonia intracellularis: use in phylogeny."  
RL Microbiology 144:0-0(0).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Dale C.J.H.;  
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Prevents misfolding and promotes the refolding and  
proper assembly of unfolded polypeptides generated under stress  
conditions (By similarity).  
CC -!- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of  
7 subunits (By similarity).  
CC -!- SIMILARITY: Belongs to the chaperonin (HSP60) family.  
DR EMBL; U45241; AAC36500.1; -.  
DR HSSP; P06139; IGR5.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0051082; F:unfolded protein binding; IEA.  
DR GO; GO:0006457; P:protein folding; IEA.  
DR InterPro; IPR001844; Chaprinin\_Cpn60.  
DR InterPro; IPR008950; GroEL-ATPase.  
DR Pfam; PF00118; Cpn60\_TCP1; 1.  
DR PRINTS; PR00298; CHAPERONIN60.  
DR PRINTS; PR0304; TCOMPLEXTCP1.  
DR PROSITE; PS00296; CHAPERONIN60; 1.  
KW ATP-binding; Chaperone; Plasmid.  
SQ SEQUENCE 548 AA; 58605 MW; 6388C431E663E498 CRC64;

Query Match 100.0%; Score 2708; DB 2; Length 548;  
Best Local Similarity 100.0%; Pred. No. 8.7e-120;  
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MASKEILFDKAREKLSRGVDKLANAVKVLGPKGRNVWIEKSPGSPVITKDGVSVAKEI	60
DB	1	MASKEILFDKAREKLSRGVDKLANAVKVLGPKGRNVWIEKSPGSPVITKDGVSVAKEI	60

```

Qy 61 ELEDKFNMGQAQVKEVAPKTSIDAGDGTATTATVLAQAIYREGVKLVAAAGNPNMAIKRGI 120
Db 61 ELEDKFNMGQAQVKEVAPKTSIDAGDGTATTATVLAQAIYREGVKLVAAAGNPNMAIKRGI 120

Qy 121 DKAVAVTKELSDITKPTRDQKEIAQVGTISANSDDTTIGNIIAAMAKVKGKGVITVVEEA 180
Db 121 DKAVAVTKELSDITKPTRDQKEIAQVGTISANSDDTTIGNIIAAMAKVKGKGVITVVEEA 180

Qy 181 KGLETTLDVVVEGKMFDRGYLSPYFVTNPEKMWCELDNPNYILCNEKKITSMKMDLPVLEQV 240
Db 181 KGLETTLDVVVEGKMFDRGYLSPYFVTNPEKMWCELDNPNYILCNEKKITSMKMDLPVLEQV 240

Qy 241 AKVNRPLIIIAEDVVEGALATLVNKLRGALQVAVKAPGFGERRKAMLEDIAILLTGGEA 300
Db 241 AKVNRPLIIIAEDVVEGALATLVNKLRGALQVAVKAPGFGERRKAMLEDIAILLTGGEA 300

Qy 301 IFEDRGIKLENVSLSSIGTAKRVVVDKENTTIVDGAGKSEDIKARVKQIRAEETSSDY 360
Db 301 IFEDRGIKLENVSLSSIGTAKRVVVDKENTTIVDGAGKSEDIKARVKQIRAEETSSDY 360

Qy 361 DREKLOERLAKLVGGVAVIHVGAATETEMKEKKDRVEDALNATRAAEEGIVPGGGTAFV 420
Db 361 DREKLOERLAKLVGGVAVIHVGAATETEMKEKKDRVEDALNATRAAEEGIVPGGGTAFV 420

Qy 421 RSIKVLDDIKPADDELALAGNIIRRSLEELPLQIAANAGYEGSIVVEKVRPKDGGFGFNA 480
Db 421 RSIKVLDDIKPADDELALAGNIIRRSLEELPLQIAANAGYEGSIVVEKVRPKDGGFGFNA 480

Qy 481 ASGEYEDLIKAGVIDPKKVTRIALQNAASVASLLLTTECAIAEKPEPKDMPMPGGMG 540
Db 481 ASGEYEDLIKAGVIDPKKVTRIALQNAASVASLLLTTECAIAEKPEPKDMPMPGGMG 540

Qy 541 MGGMDGMY 548
Db 541 MGGMDGMY 548

```

## RESULT 2

```

Q72AL6 Q72AL6 PRELIMINARY; PRT; 547 AA.
AC Q72AL6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Chaperonin, 60 kDa.
GN NamesgroEL; OrderedLocusNames=DVU1976;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]
SEQUENCE FROM N.A.
RX PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.F., Sehadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA Lonhart S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Daviden T.M., Zafar N., Zhou L., Radune D.,
RA Dmitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
RL Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
CC -1- FUNCTION: Prevents unfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
DR EMBL; AE017315; AAS96452.1; -.
DR HSP; F06139; 1kD.
DR TIGR; DVU1976; -.

```

```

DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0051082; F:unfolded protein binding; IEA.
DR GO; GO:0006457; P:protein folding; IEA.
DR InterPro; IPR001844; Chaprinin_Cpn60.
DR InterPro; IPR002423; Cpn60/TCP-1.
DR InterPro; IPR008950; GroEL-ATPase.
DR Pfam; PF00118; Cpn60_TCP1; 1.
DR PRINTS; PR00298; CHAPERONIN60.
DR PRINTS; PR00304; TCOMPLEXTCP1.
DR PROSITE; PS00296; CHAPERONIN_CPN60; 1.
KW ATP-binding; Chaperone; Complete proteome.
SQ SEQUENCE 547 AA; 58389 MW; 5743A7CAD92C529A CRC64;

Query Match 90.9%; Score 2462.5; DB 2; Length 547;
Best Local Similarity 89.2%; Pred. No. 3.1e-108;
Matches 489; Conservative 36; Mismatches 22; Indels 1; Gaps 1;

Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVLGPKGRNVVIEKSFSGSVITKDGVSVAKEI 60
Db 1 MASKEILFDTKAREKLSRGVDKLANAVKVLGPKGRNVVIEKSFSGSVITKDGVSVAKEI 60

Qy 61 ELEDKFNMGQAQVKEVAPKTSIDAGDGTATTATVLAQAIYREGVKLVAAAGNPNMAIKRGI 120
Db 61 ELEDKFNMGQAQVKEVAPKTSIDAGDGTATTATVLAQAIYREGVKLVAAAGNPNMAIKRGI 120

Qy 121 DKAVAVTKELSDITKPTRDQKEIAQVGTISANSDDTTIGNIIAAMAKVKGKGVITVVEEA 180
Db 121 DKAVAVTKELSDITKPTRDQKEIAQVGTISANSDDTTIGNIIAAMAKVKGKGVITVVEEA 180

Qy 181 KGLETTLDVVVEGKMFDRGYLSPYFVTNPEKMWCELDNPNYILCNEKKITSMKMDLPVLEQV 240
Db 181 KGLETTLDVVVEGKMFDRGYLSPYFVTNPEKMWCELDNPNYILCNEKKITSMKMDLPVLEQV 240

Qy 241 AKVNRPLIIIAEDVVEGALATLVNKLRGALQVAVKAPGFGERRKAMLEDIAILLTGGEA 300
Db 241 AKVNRPLIIIAEDVVEGALATLVNKLRGALQVAVKAPGFGERRKAMLEDIAILLTGGEA 300

Qy 301 IFEDRGIKLENVSLSSIGTAKRVVVDKENTTIVDGAGKSEDIKARVKQIRAEETSSDY 360
Db 301 IFEDRGIKLENVSLSSIGTAKRVVVDKENTTIVDGAGKSEDIKARVKQIRAEETSSDY 360

Qy 361 DREKLOERLAKLVGGVAVIHVGAATETEMKEKKDRVEDALNATRAAEEGIVPGGGTAFV 420
Db 361 DREKLOERLAKLVGGVAVIHVGAATETEMKEKKDRVEDALNATRAAEEGIVPGGGTAFV 420

Qy 421 RSIKVLDDIKPADDELALAGNIIRRSLEELPLQIAANAGYEGSIVVEKVRPKDGGFGFNA 480
Db 421 RSIKVLDDIKPADDELALAGNIIRRSLEELPLQIAANAGYEGSIVVEKVRPKDGGFGFNA 480

Qy 481 ASGEYEDLIKAGVIDPKKVTRIALQNAASVASLLLTTECAIAEKPEPKDMPMPGGMG 540
Db 481 ASGEYEDLIKAGVIDPKKVTRIALQNAASVASLLLTTECAIAEKPEPKDMPMPGGMG 540

Qy 541 MGGMDGMY 548
Db 541 MGGMDGMY 548

RESULT 3
Q747C7 Q747C7 PRELIMINARY; PRT; 544 AA.
AC Q747C7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE 60 kDa chaperonin.
GN Name=groEL; OrderedLocusNames=GSU3340;
OS Geobacter sulfurreducens.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC Geobacteraceae; Geobacter.
OX NCBI_TaxID=35554;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN=PCA / ATCC 51573;  
RX PubMed=14671304; DOI=10.1126/science.1088727;  
RA Mathe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,  
RA Heidelberg J.F., Wu D., Ward N.L., Beanan M.J., Dodson R.J.,  
RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,  
RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,  
RA Daviden T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,  
RA Weidman J.F., Khouri H.M., Feldblyum T.V., Utterback T.R.,  
RA Van Aken S.E., Lovley D.R., Fraser C.M.;  
RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface  
RT environments.";  
RL Science 302:1967-1969(2003).  
CC -1- FUNCTION: Prevents misfolding and promotes the refolding and  
CC proper assembly of unfolded polypeptides generated under stress  
CC conditions (By similarity).  
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of  
CC 7 subunits (By similarity).  
CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.  
DR EMBL; AE017180; AAR36730.1; -.  
DR HSP; P06139; IKID.  
DR TIGR; GSU3340; -.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0051082; F:unfolded protein binding; IEA.  
DR GO; GO:0006457; P:protein folding; IEA.  
DR InterPro; IPR001844; Chaprnl\_Cpn60.  
DR InterPro; IPR002423; Cpn60/TCP-1.  
DR InterPro; IPR008950; GroEL-ATPase.  
DR Pfam; PF00118; Cpn60\_TCP1; 1.  
DR PRINTS; PR00298; CHAPERONIN60.  
DR PRINTS; PR00304; TCOMPLEXTCP1.  
DR PROSITE; PS00296; CHAPERONIN\_CPN60; 1.  
KW ATP-binding; Chaperone; Complete proteome.  
SQ SEQUENCE 544 AA; 58148 MW; 5266B4E9126B4A98 CRC64;  
  
Query Match 70.7%; Score 1967.5; DB 2; Length 544;  
Best Local Similarity 71.9%; Pred. No. 6.3e-85;  
Matches 391; Conservative 70; Mismatches 82; Indels 1; Gaps 1;  
  
Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSPGVITKDGVSVAKEI 60  
Db 1 MAARIKFDQEGRNAILKGVNTLADAVKVTLPKGRNVVIEKSPGVITKDGVSVAKEI 60  
  
Qy 61 ELEDKFNMGQAQVKEVAPKTSIDAGDGTATTATVLAQAIYREGVKLVAAAGNPMIAIKGI 120  
Db 61 ELEDKFNMGQAQVKEVASKTSIDVAGDGTATTATVLAQAIYRQSKLVAAAGNPMIAIKGI 120  
  
Qy 121 DKAVAVVTKELSDITKPTDQKEIAQVGTISANSDDTTIGNIIAEAMAKVGKGGVITVEEA 180  
Db 121 DKAVETIVAEKLSISKPIKDKHEIAQVGTISANDKTIIGDIIAQMKEKVGKGGVITVEEA 180  
  
Qy 181 KGLFTLLDVVVEGMKFDGRGYSIPYFVTNPKMVCLENDNPIYLCNEKKTITSMKMDLPVLEQV 240  
Db 181 KAMETSLVETVEGMQFDRGYSIPYFVTNPKMVCLENDNPIYLCNEKKTITSMKMDLPVLEQV 240  
  
Qy 241 AKVNRPLIIIAEDVEGALATLVNKLRGALQVAVKAPGEGRRKAMLEDIAITLTGGEA 300  
Db 241 AKSGRPLIIIAEDVEGALATLVNKLRLGVNLICAVKAPGEGRRKAMLEDIAITLTGQV 300  
  
Qy 301 IFDRGIKLENVSLSSIGTAKRVVVDKENTTIVDGACKSEDIKARVKIQAQIETSSDY 360  
Db 301 ISEIGNKLENTTMDMLGRKRIYVDKNTIIDGKEADIQRVKIQIAQIETSSDY 360  
  
Qy 361 DREKLQRLAKLVGGVAVIHVGAATETEMKEKRDVEDALNATRAAEEVPGGGTAFV 420  
Db 361 DREKLQRLAKLVGGVAVIKVGAATETEMKEKARVEDALHATRAAEEVPGGGVAVL 420  
  
Qy 421 RSIKVLDDIKPADDDDELAGNIIPRSLEELPQIQAANAGVEGSIWVEKPEKPGGFNA 480  
Db 421 RALASLDALSIFTEQQF-GVNVIKRSLEELPQIQAANAGVDGSIWVDKVNKSQDAFGNA 479  
  
Qy 481 ASGEYEDILIKAGVIDPKKVTRIALQNAASVASLLETTTECAIAEKPEPKDMPGCGMG 540  
Db 480 ADEYVDMLAAGIIDPTKVSRSALQNAASVAGLMLTTEMIADPKPEAPKMPGCGMG 539

Qy 541 MGGM 544  
Db 540 MGGM 543  
  
RESULT 4  
Q6NOB9  
ID Q6NOB9 PRELIMINARY; PRT; 547 AA.  
AC Q6NOB9;  
DT 05-JUL-2004 (T-EMBLrel. 27, Created)  
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)  
DT 08-JUL-2004 (T-EMBLrel. 27, Last annotation update)  
DE GROEL.  
GN ORFNames=orf4708;  
OS Magnetospirillum gryphiswaldense.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;  
OC Rhodospirillaceae; Magnetospirillum.  
OX NCBI\_TaxID=55518;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gruenberg K., Mueller E.C., Otto A., Reszka R., Linder D., Kube M.,  
RA Reinhardt R., Schueler D.;  
RT "Biochemical and proteomic analysis of the magnetosome membrane in  
RT Magnetospirillum gryphiswaldense.";  
RL Appl. Environ. Microbiol. 70:1040-1050(2004).  
CC -1- FUNCTION: Prevents misfolding and promotes the refolding and  
CC proper assembly of unfolded polypeptides generated under stress  
CC conditions (By similarity).  
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of  
CC 7 subunits (By similarity).  
CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.  
DR EMBL; BX640510; CAB45331.1; -.  
DR HSP; P06139; IKID.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0051082; F:unfolded protein binding; IEA.  
DR GO; GO:0006457; P:protein folding; IEA.  
DR InterPro; IPR001844; Chaprnl\_Cpn60.  
DR InterPro; IPR002423; Cpn60/TCP-1.  
DR InterPro; IPR008950; GroEL-ATPase.  
DR Pfam; PF00118; Cpn60\_TCP1; 1.  
DR PRINTS; PR00298; CHAPERONIN60.  
DR PRINTS; PR00304; TCOMPLEXTCP1.  
DR PROSITE; PS00296; CHAPERONIN\_CPN60; 1.  
KW ATP-binding; Chaperone.  
SQ SEQUENCE 547 AA; 57690 MW; BCC03A64D0AC2AE4 CRC64;

Query Match 70.2%; Score 1900.5; DB 2; Length 547;  
Best Local Similarity 71.1%; Pred. No. 9e-82;  
Matches 388; Conservative 60; Mismatches 97; Indels 1; Gaps 1;  
  
Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSPGVITKDGVSVAKEI 60  
Db 1 MAAKEVKFSTEARAKMLRGVDILADAVKVTLPKGRNVVIEKSPGVITKDGVSVAKEI 60  
  
Qy 61 ELEDKFNMGQAQVKEVAPKTSIDAGDGTATTATVLAQAIYREGVKLVAAAGNPMIAIKGI 120  
Db 61 ELEDKFNMGQAQVREVASKTSIDLVDGDTATTATVLAQAIYREGKNAAGNPMIAIKGI 120  
  
Qy 121 DKAVAVVTKELSDITKPTDQKEIAQVGTISANSDDTTIGNIIAEAMAKVGKGGVITVEEA 180  
Db 121 DLAVEAVADVVKGRKVKSTNEEIAQVGTISANGERDIGAKIAEAMSKVNEGVTVEEA 180  
  
Qy 181 KGLFTLLDVVVEGMKFDGRGYSIPYFVTNPKMVCLENDNPIYLCNEKKTITSMKMDLPVLEQV 240  
Db 181 KGFETELDVVVEGMQFDRGYSIPYFVTNPKMVCLENDNPIYLCNEKKTITSMKMDLPVLEAV 240  
  
Qy 241 AKVNRPLIIIAEDVEGALATLVNKLRGALQVAVKAPGEGRRKAMLEDIAITLTGGEA 300  
Db 241 VQSGRPLIIIAEDVEGALATLVNKLRLGVNLICAVKAPGEGRRKAMLEDIAITLTGQV 300  
  
Qy 301 IFDRGIKLENVSLSSIGTAKRVVVDKENTTIVDGACKSEDIKARVKIQAQIETSSDY 360  
Db 301 IFDRGIKLENVSLSSIGTAKRVVVDKENTTIVDGACKSEDIKARVKIQAQIETSSDY 360

```

Db 301 ISEDLGKLESVLEMLGTAKRVTTTKEDTTIVDAGARGDIEARCKQIRAQIEETTSY 360
Qy 361 DREKQERLAKLVGGVAVIHVGATETEEMKEDRVEDALNATRAAAVEGIVPGGTAFFV 420
Db 361 DREKQERLAKLVGGVAVIHVGATETEEMKEDRVEDALNATRAAAVEGIVPGGTAFFV 420
Qy 421 RSIKVLDDIKPADDELALGNIIRRSLEELRQIAANAGYEGSVIVKREPKD-GFGFN 479
Db 421 YSVRALDGIADVNDQKVIDIVRALQSPVRQIAENAGHDGAVVAGKLEATDTNFGFD 480
Qy 480 AAGEYEDLIKAGVIDPKKVTIRIALQNAASVASLLLTTECAIAEKPKPKMDMPGGMG 539
Db 481 AQTGYVDMIKAGIIDPVKVVRTALQDAASVAGLLITTEAMIAEKKPKKDGSPSPGGMG 540
Qy 540 GMGMD 545
Db 541 GMGMD 546

RESULT 5
CH60_RALSO STANDARD; PRT; 547 AA.
ID CH60_RALSO STANDARD; PRT; 547 AA.
AC Q8Y1P8;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
GN Names: groEL, mopA; OrderedLocNames=RSC0642;
GN ORFNames=RS01546;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RC MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan N., Robert C., Saurin W., Schlex T.,
RA Siguer P., Thebaut P., Whalen M., Wincken M., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002)
CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annouce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL646060; CAD14172.1; -.
CC DR HSSP; P06139; 1GR5.
CC DR HAMAP; MF 00600; -.
CC DR InterPro; IPR001844; Chaperonin Cpn60.
CC DR InterPro; IPR002423; Cpn60/TCF-1.
CC DR InterPro; IPR008950; GroEL-ATPase.
CC DR Pfam; PF00118; Cpn60 TCF1; 1.
CC DR PRINTS; PR00298; CHAPERONIN60.
CC DR PRINTS; PR00304; TCOMPLEXTCP1.
CC DR PROSITE; PS00296; CHAPERONINS_CPN60; 1.
CC KW ATP-binding; Chaperone; Complete proteome.
CC SQ SEQUENCE 547 AA; 57404 MW; 0487B4650867C25A CRC64;

```

```

Query Match 69.8%; Score 1891.5; DB 1; Length 547;
Best Local Similarity 69.5%; Pred. No. 2.4e-81;
Matches 380; Conservative 70; Mismatches 96; Indels 1; Gaps 1;

Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVLGPKGRNVVIEKSFSGSVITKDGVSVAKEI 60
Db 1 MAAKDVVFGDAARAKWVEGVNLANAVKVLGPKGRNVVLEISFGGPTVKDGVSVAKEI 60
Qy 61 ELSDKFNMGAAQWKEVAPKTSIDAGDGTATATLAQAIYREGVKLVAAAGNPMIAIRGI 120
Db 61 ELKDKLQNMGAQWKEVASKTSIDAGDGTATATLAQSIYREGMKYVAAAGNPMIDLRGI 120
Qy 121 DKAVAVVTKELSDITKPTDRQKIAQVGTISANSDDTTIGNI IAEAMAKVGKGVITVVEA 180
Db 121 DKAVAAAEELKKISKPTTTSKIAQVGTISANSDESIGARIAEMDKVKEGVITVEDG 180
Qy 181 KGLETTLDVVVEGKMFDRGYLSPYFVTNPERKWCVELDNPYILCNEKKITSMKMDLPILQV 240
Db 181 KSLDELDDVVVEGKMFDRGYLSPYFINNPEKQVQLDNPFVLLFDKKISINRDLPLVLEQV 240
Qy 241 AKVNRPLIIIAEDVEGEALATLVNKLRGALQVAVKAPGFGERRKAMLEDIAILLTGEA 300
Db 241 AKAGRPLIIIAEDVEGEALATLVNNIRGLTKAAVKAPGFGERRKAMLEDIAILLTGGV 300
Qy 301 IFEDRGIKLENVSLSSLTGAKRVVIDKENTTIVDAGKSEDIKARVKQIRAQIEETSSY 360
Db 301 IAEVGLTLEKATLNDLGQAKRVEIKENTTIDGAGDARNI EARVKVQRAQIEEATSDY 360
Qy 361 DREKQERLAKLVGGVAVIHVGATETEEMKEDRVEDALNATRAAAVEGIVPGGTAFFV 420
Db 361 DREKQERLAKLVGGVAVIHVGATETEEMKEDRVEDALNATRAAAVEGIVAGGVALL 420
Qy 421 RSIKVLDDIKPADDELALGNIIRRSLEELRQIAANAGYEGSVIVKREPKDGFNFNA 480
Db 421 RARALISGLKANADQAGIKIVLAMEEPLRQIVTNAGDEASVVVANVITAGKNGYNA 480
Qy 481 AAGEYEDLIKAGVIDPKKVTIRIALQNAASVASLLLTTECAIAEKPEPKMDMPGGMG 540
Db 481 STGEYGDVEMGVLDPTKVTIRIALQNAASVASLLMTTDCAVAEIPKDDAAMP-PG 539
Qy 541 GMGMD 547
Db 541 GMGMD 546

RESULT 6
CH60_ACEAC STANDARD; PRT; 546 AA.
ID CH60_ACEAC STANDARD; PRT; 546 AA.
AC Q8GBD2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
GN Name: groL; Synonyms=groEL;
OS Acetobacter aceti.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Acetobacteraceae; Acetobacter.
OX NCBI_TaxID=435;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 3283;
RC Okamoto-Kainuma A., Yan W., Kadono S., Tayama K., Koizumi Y.,
RC Yanagida F.;
RT "Cloning and characterization of groEL operon in Acetobacter aceti.";
RL J. Biosci. Bioeng. 94:140-147(2002).
CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.

```



```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; AB081596; BAC16232.1; -.
CC HAMAP; MF 00600; -. 1.
CC InterPro; IPR001844; Chaprinin_Cpn60.
CC InterPro; IPR002423; Cpn60/TCP-1.
CC Pfam; PF00118; Cpn60 TCP1; 1.
CC PRINTS; PR00298; CHAPERONIN60.
CC PRINTS; PR00304; TCOMPLEXTCP1.
CC PROSITE; PS00296; CHAPERONINS_CPN60; 1.
KW ATP-binding; Chaperone.
SQ SEQUENCE 546 AA; 58106 MW; 38476AF42F2CD5CC CRC64;

Query Match          69.4%; Score 1880; DB 1; Length 546;
Best Local Similarity 69.8%; Pred. No. 8.3e-81;
Matches 381; Conservative 69; Mismatches 94; Indels 2; Gaps 2;

Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSPGSPVITKDGVSVAKEI 60
Db 1 MAADKVFQADARQRMAGVLDLADAVKVTLPKGRNVVIEKSPGSPVITKDGVSVAKEI 60

Qy 61 ELEDKFNMGCAQWKEVAPKTSIAGDGTATTATVLAQAIYREGVKLVAAAGNPMIAIKRGI 120
Db 61 ELADKFNMGCAQWKEVAPKTSIAGDGTATTATVLAQAIYREGVKLVAAAGNPMIAIKRGI 120

Qy 121 DKAVAVVTKELSDITKTPDQKEIAQVGTISANSDDTTIGNIIAEAMAKVKGKGGVITVEEA 180
Db 121 DKAVAVVIEELKQKAVKVTTPAETQAQVGTISANGSEIGQMSAMQKVGSEGVITVEEA 180

Qy 181 KGLTELLDVVEGKMFDRGYLSPYFVTNPKMVCLELDNPYILCNEKKTITSMKMDLPILAEQV 240
Db 181 KHFTELLDVVEGKMFDRGYLSPYFVTNPKMTADLENPYILIEKLSLQPLLESV 240

Qy 241 AKNRPILLIIAEVDGEGALATLVNKLKRGALQVAVKAPGFGERRKAMLEDAIALLTGGEA 300
Db 241 VQSGRPLIIIAEDVDGEGALATLVNKLKRGALQVAVKAPGFGERRKAMLEDAIALLTGGEA 300

Qy 301 IFEDRGKLENVSLSSIGTAKRVVIDKENTTIIVDGAGKSEDIKARVKQIRAOIETSSDY 360
Db 301 ISEDLGKLETVLNLMLGTAKRVVIDKENTTIIVDGAGKADDIKGRVKQIRAOIETSSDY 360

Qy 361 DREKQLERLAKLVGGVAVIHVGAATETEMKCKDRVEDALNATRAAAVEEGIVPGGGTAFV 420
Db 361 DREKQLERLAKLVGGVAVIRVGGSTEVEKVKERKDRVDALHATRAAAVEEGIVPGGGTALA 420

Qy 421 RSKVLDDIIPADDELALGNIIRRSLEELPQRTAANAGVEGSIIVSKVRPKDGRGPN 479
Db 421 RATLKLKELGLHYNDQVGGDIIRALQAPLRQAHNAGEDGAVIANKVLNSDYNGFD 480

Qy 480 AASGEYEDLIKAGVIDPKKVTTRIALQNAASVASLTLTTTECAIAEKPKPKMPGGMG 539
Db 481 AQAGEYKLVVAGIIDPAKVVRALQNAASVAGLLITTEAMVAPERPE-KKAAPAGGPDGMG 539

Qy 540 GMGGMD 545
Db 540 GMGGMD 545

RESULT 7
ID Q75T66 PRELIMINARY; PRT; 547 AA.
AC Q75T66;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

```

```

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Molecular chaperone GroEL.
GN Name-groELb;
OS Burkholderia pickettii (Pseudomonas pickettii).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
ON NCBI_TaxID=329;
RX SEQUENCE FROM N.A.
RA Tokumaru N., Takizawa N., Kiyohara H.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -!- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -!- SIMILARITY: Belongs to the chaperonin (HSP60) family.
DR EMBL; AB126866; BAD06928.1; -.
DR HSP; P06139; IKID.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0051082; F:unfolded protein binding; IEA.
DR GO; GO:0006457; P:protein folding; IEA.
DR InterPro; IPR001844; Chaprinin_Cpn60.
DR InterPro; IPR002423; Cpn60/TCP-1.
DR InterPro; IPR008950; GroEL-ATPase.
DR InterPro; IPR000215; Prot_inh_serpin.
DR Pfam; PF00118; Cpn60 TCP1; 1.
DR PRINTS; PR00298; CHAPERONIN60.
DR PRINTS; PR00304; TCOMPLEXTCP1.
DR PROSITE; PS00296; CHAPERONINS_CPN60; 1.
DR PROSITE; PS00284; SERPIN; UNKNOWN_1.
KW ATP-binding; Chaperone.
SQ SEQUENCE 547 AA; 57319 MW; 089B3FD1B726C080 CRC64;

Query Match          69.4%; Score 1878.5; DB 2; Length 547;
Best Local Similarity 68.9%; Pred. No. 9.7e-81;
Matches 377; Conservative 69; Mismatches 100; Indels 1; Gaps 1;

Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSPGSPVITKDGVSVAKEI 60
Db 1 MAADKVFQADARQRMAGVLDLADAVKVTLPKGRNVVIEKSPGSPVITKDGVSVAKEI 60

Qy 61 ELEDKFNMGCAQWKEVAPKTSIAGDGTATTATVLAQAIYREGVKLVAAAGNPMIAIKRGI 120
Db 61 ELKDKLQNMGAQWKEVAPKTSIAGDGTATTATVLAQAIYREGVKLVAAAGNPMIAIKRGI 120

Qy 121 DKAVAVVTKELSDITKTPDQKEIAQVGTISANSDDTTIGNIIAEAMAKVKGKGGVITVEEA 180
Db 121 DKAVTAAVESLKKVKSPFTTSKEIAQVGAISANSDDTSIGERIAEAMDKVKGEGVITVEDG 180

Qy 181 KGLTELLDVVEGKMFDRGYLSPYFVTNPKMVCLELDNPYILCNEKKTITSMKMDLPILAEQV 240
Db 181 KSLADELVEVGEQFDRGYLSPYFINNPEKQVQLDSPFVLLPDKKVSINRDLPLVLEQV 240

Qy 241 AKNRPILLIIAEVDGEGALATLVNKLKRGALQVAVKAPGFGERRKAMLEDAIALLTGGEA 300
Db 241 AKAGRPLIIIAEDVDGEGALATLVNKNIRGILKTAAVKAPGFGDRRKAMLEDAIALLTGTV 300

Qy 301 IFEDRGKLENVSLSSIGTAKRVVIDKENTTIIVDGAGKSEDIKARVKQIRAOIETSSDY 360
Db 301 IAEIIGLTKAGLNDLQQAQRIEIGKENTIIIDGADGAGIENGVRQIRAOIETSSDY 360

Qy 361 DREKQLERLAKLVGGVAVIHVGAATETEMKCKDRVEDALNATRAAAVEEGIVPGGGTAFV 420
Db 361 DREKQLERLAKLVGGVAVIKVGAATEVEKVKERKARVEDALHATRAAAVEEGIVPGGVALL 420

Qy 421 RSKVLDDIIPADDELALGNIIRRSLEELPQRTAANAGVEGSIIVSKVRPKDGRGPN 480
Db 421 RARAASALTGENADQAGIKIVLRAMEEPLQVILNAGEASVVAKVIVGKNGYGYNA 480

Qy 481 AASGEYEDLIKAGVIDPKKVTTRIALQNAASVASLTLTTTECAIAEKPKPKMPGGMG 540
Db 481 ASGEYGLVEMGVLDPTKVTTRIALQNAASVASLTLTTTDCVAESPKDESAPAMP-GMGG 539

```

```
QY 541 MGGMDGM 547
Db 540 MGGMEGM 546

RESULT 8
Q6ZCZ7 PRELIMINARY; PRT; 547 AA.
AC Q6ZCZ7
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Chaperonin GroEL.
GN Name=GroEL;
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cyctobacterineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN [1]
RP SEQUENCE FROM N.A.
RA Carrero-Lerida J., Moraleda-Munoz A., Perez J., Hofmann D.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Munoz-Dorado J.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
DR HSP60; P06139; IAOB.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0051082; F:unfolded protein binding; IEA.
DR GO; GO:0006457; P:protein folding; IEA.
DR InterPro; IPR001844; Chaperin Cpn60.
DR InterPro; IPR002423; Cpn60/TCF-1.
DR InterPro; IPR008950; GroEL-ATPase.
DR Pfam; PF00118; Cpn60_TCF1; 1.
DR PRINTS; PR00298; CHAPERONIN60.
DR PROSITE; PS00304; TCOMPLEXTCP1.
DR PROSITE; PS00296; CHAPERONINS_CPN60; 1.
KW ATP-binding; Chaperone.
SQ SEQUENCE 547 AA; 58145 MW; 8884711C95EACB0F CRC64;

Query Match 69.2%; Score 1874.5; DB 2; Length 547;
Best Local Similarity 69.2%; Pred. No. 1.5e-80;
Matches 379; Conservative 69; Mismatches 99; Indels 1; Gaps 1;

QY 1 MASKELFDKAREKLSRGVDKLANAVKVLGPKGRNVVIEKSPGSPVITKDGVSVAKEI 60
Db 1 MAAKEIFPHOSAREAILRGVRLSDAVAVTLGPKGRNVVIEKSPGSPVITKDGVTVAKEI 60

QY 61 ELEDKENMGAQVKEVAPKTSIDAGGTATVLAQAIYREGVKLVAAAGNPMAIKRGI 120
Db 61 DLENKFNMGQAQVKEVASKTSKAGGDTTATVLAQAIYEEGLKLVAAAGHSPMDLKRG 120

QY 121 DKAVAVVTKELSDITKPTDQKEIAQVGTISANSDDTTIGNIIAEMAKVGGVITVEEA 180
Db 121 DKAVEVVVVELKSLSKPTADKKAITQVGTISANGDETIGAIADAMEKVGKGVITVEEA 180

QY 181 KGLETLDDVVEGKMFDRGYLSPFTVTPNPKVCELDNPNYILCNKKITSMKMDLPLEQV 240
Db 181 KGLETLDDVVEGKMFDRGYLSPFTVTPNPKVCELDNPNYILCNKKITSMKMDLPLEQV 240

QY 241 AKVNRPLLIIEADVEGEGALATLVNKLRLGALQVAVKAPGGRKRWKLELDIALLTGEEA 300
Db 241 ARSGKPLIIADIEGEGALATLVNKLRLGALQVAVKAPGGRKRWKLELDIALLTGEEA 300

QY 301 IFEDRGIKLENVSLSSLGTAQVVKIDKENTTIYDGAQKSEDIKARVKQIRAQIEETSSDY 360
Db 301 VSEDLGHKFETLTLDLGRAKRVTVDKNTTVVDGVTGKAAIEGRIKLTQIDSVTSDY 360

QY 361 DREKLOERLAKLVGGVAVIHVGAATETEMKEKKORVEDALNATRAAEVGGGTAFV 420
Db 361 DREKLOERLAKLVGGVAVIHVGAATETEMKEKKARVEDALHATRAAEVGGGVAYL 420

QY 421 RSIKVLDDIKPADDELALGNIIRRSLEELPRLQIAANAGVEGSIIVVEKPKDGFQFNA 480
Db 421 RALPALEKLPKGGEQDF-GVAIRRALQELPKRIASNAGVAVINKVRIGTGAFGNA 479

QY 481 ASGEYEDLIKAGVIDPKKVTRIALQNAASVASLLLTTECAIAEKPEPKOMPMPGGMG 540
Db 480 RTEVYEDLEKAGVIDPKKVERTALQNAASVASLLLTTEAMVAERPKKKGAGAGMPD 539

QY 541 MGGMDGM 548
Db 540 YGGDDMDY 547

RESULT 9
CH60_BURCE
ID CH60_BURCE STANDARD; PRT; 546 AA.
AC Q9ZFEO;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (GroEL protein).
GN Name=groL; Synonyms=groEL, mopA;
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NTC 10744;
RA Zysk G., Splettoesser W.D., Neubauer H.;
RT "Nucleotide sequence comparison of the groB operon of Burkholderia
RT spp.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Prevents misfolding and promotes the refolding and
CC proper assembly of unfolded polypeptides generated under stress
CC conditions (By similarity).
CC -1- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of
CC 7 subunits (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the chaperonin (HSP60) family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF104907; AAC79087.1; -.
CC HSP60; P06139; ICR5.
CC HAMAP; MF_00600; -. 1.
CC InterPro; IPR001844; Chaperin Cpn60.
CC InterPro; IPR002423; Cpn60/TCF-1.
CC InterPro; IPR008950; GroEL-ATPase.
CC Pfam; PF00118; Cpn60_TCF1; 1.
CC PRINTS; PR00298; CHAPERONIN60.
CC PRINTS; PR00304; TCOMPLEXTCP1.
CC PROSITE; PS00296; CHAPERONINS_CPN60; 1.
CC ATP-binding; Chaperone.
SQ SEQUENCE 546 AA; 56980 MW; 174B9934345E7315 CRC64;

Query Match 69.1%; Score 1872.5; DB 1; Length 546;
Best Local Similarity 68.4%; Pred. No. 1.9e-80;
Matches 374; Conservative 75; Mismatches 95; Indels 3; Gaps 2;
```



Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
Burkholderiaceae; Burkholderia.  
NCBI\_TaxID=28450;  
[1]  
SEQUENCE FROM N.A.  
Woo P.C.Y., Leung P.K.L.;  
Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
-!- FUNCTION: Prevents misfolding and promotes the refolding and  
proper assembly of unfolded polypeptides generated under stress  
conditions (By similarity).  
-!- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of  
7 subunits (By similarity).  
-!- SIMILARITY: Belongs to the chaperonin (HSP60) family.  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
EMBL; AF287633; AAG32927.1; -;  
HSSP; P06139; 1GR5.  
HAMAP; MF\_00600; -; 1.  
InterPro; IPR001844; Chaprin Cpn60.  
InterPro; IPR002423; Cpn60/TCP-1.  
InterPro; IPR008950; GroEL-ATPase.  
Pfam; PF00118; Cpn60\_TCP1; 1.  
PRINTS; PR00298; CHAPERONIN60.  
PRINTS; PR00304; TCOMPLEXTCP1.  
PROSITE; PS00296; CHAPERONINS\_CPN60; 1.  
ATP-binding; Chaperone.  
SEQUENCE 546 AA; 57145 MW; E9E9366EBBDB6ADB CRC64;  
-----  
Query Match 69.0%; Score 1868.5; DB 1; Length 546;  
Best Local Similarity 68.2%; Pred. No. 2.9e-80;  
Matches 373; Conservative 75; Mismatches 96; Indels 3; Gaps 2;  
QY 1 MASKEILFPAKAREKLSRGVDKLANAVKVTGLPGKRNVIIEKSGFSPVITKDGVSVAKEI 60  
DB 1 MAAKDVVFGDSARAKMVEGVNLANAVKVTGLPGKRNVLERSFGPGPTVTKDGVSVAKEI 60  
QY 61 ELEDKPENNGAQWKEVAPKTSIDAGDGTGTTATVLAQAIYREGVKLVAAGRNPMIAIKRGI 120  
DB 61 ELKDXLQNNGAQWKEVASKTSIDNAGDGTGTTATVLAQISVREGMKYVASGMNPMDLKRG 120  
QY 121 DKAVAVTKELSDITKPTDQKEIAQVGTISANSDDTTIGNIIAEAMAKVKGKGVITVVEA 180  
DB 121 DKAVAAVEELKKISKPTCTTNKEIAQVGSISANSDDTTIGNIIAEAMAKVKGKGVITVVEA 180  
QY 481 ASGEYEDLIKAGVIDPKKVTYTRIALQNAASVASLLLTTECAIAEAKPEPKKMPMPGGMG 540  
-----  
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
Burkholderiaceae; Burkholderia.  
NCBI\_TaxID=28450;  
[1]  
SEQUENCE FROM N.A.  
Woo P.C.Y., Leung P.K.L.;  
Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
-!- FUNCTION: Prevents misfolding and promotes the refolding and  
proper assembly of unfolded polypeptides generated under stress  
conditions (By similarity).  
-!- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of  
7 subunits (By similarity).  
-!- SIMILARITY: Belongs to the chaperonin (HSP60) family.  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
EMBL; AF287633; AAG32927.1; -;  
HSSP; P06139; 1GR5.  
HAMAP; MF\_00600; -; 1.  
InterPro; IPR001844; Chaprin Cpn60.  
InterPro; IPR002423; Cpn60/TCP-1.  
InterPro; IPR008950; GroEL-ATPase.  
Pfam; PF00118; Cpn60\_TCP1; 1.  
PRINTS; PR00298; CHAPERONIN60.  
PRINTS; PR00304; TCOMPLEXTCP1.  
PROSITE; PS00296; CHAPERONINS\_CPN60; 1.  
ATP-binding; Chaperone.  
SEQUENCE 546 AA; 57145 MW; E9E9366EBBDB6ADB CRC64;  
-----  
Query Match 69.0%; Score 1868.5; DB 1; Length 546;  
Best Local Similarity 68.2%; Pred. No. 2.9e-80;  
Matches 373; Conservative 75; Mismatches 96; Indels 3; Gaps 2;  
QY 1 MASKEILFPAKAREKLSRGVDKLANAVKVTGLPGKRNVIIEKSGFSPVITKDGVSVAKEI 60  
DB 1 MAAKDVVFGDSARAKMVEGVNLANAVKVTGLPGKRNVLERSFGPGPTVTKDGVSVAKEI 60  
QY 61 ELEDKPENNGAQWKEVAPKTSIDAGDGTGTTATVLAQAIYREGVKLVAAGRNPMIAIKRGI 120  
DB 61 ELKDXLQNNGAQWKEVASKTSIDNAGDGTGTTATVLAQISVREGMKYVASGMNPMDLKRG 120  
QY 121 DKAVAVTKELSDITKPTDQKEIAQVGTISANSDDTTIGNIIAEAMAKVKGKGVITVVEA 180  
DB 121 DKAVAAVEELKKISKPTCTTNKEIAQVGSISANSDDTTIGNIIAEAMAKVKGKGVITVVEA 180  
QY 481 ASGEYEDLIKAGVIDPKKVTYTRIALQNAASVASLLLTTECAIAEAKPEPKKMPMPGGMG 540  
-----

```

Qy 181 KGLTTLDVVGKMPDRGYLSPYFVTNPKVWCVDLNPYILCNKKITSMKMDLPILEQV 240
Db 181 KSLADELDVVGKMPDRGYLSPYFVTNPKVQVAVLDNPFVLLHDKVSNIRDLPLVLEQV 240

Qy 241 AKVNRPLLIITAEVDVEGALATLVNKLKRGALQVAVKAPGGERKAWLEDIAITLTGGEA 300
Db 241 AKAGRPLLIITAEVDVEGALATLVNNGIRGLKTVAVKAPGGERKAWLEDIAITLTGGQV 300

Qy 301 IFEDGKLENVSLSSLTAKRVVIDKENTTIVDGAGKSEDIKARVKQIRAIQIBETSSDY 360
Db 301 IAEETGLTLEKATLAEGLQAKRIEVRKENTTIIDGAGEAASIEARVKQIRQIEEATSDY 360

Qy 361 DREKQERLAKLVGGVAVIHVGAATETEMKEDKDRVEDALNATRAAEEGIVPGGGTAFV 420
Db 361 DREKQERLAKLVGGVAVIKVGAATEVEMKEKARVEDALHATRAAEEGIVAGGGVALI 420

Qy 421 RSIKVLDDIKPADDELALGNIIRRSLEELPQIRQIAANAGVEGSIIVKVRPKDGFNFNA 480
Db 421 RARTAIAGLTGADQAGIKIVIRAMEEPLRQIVITNGGEEASVVVAAVAAGKNYGINA 480

Qy 481 ASGEYEDLIKAGVIDPKKVTTRIALQNAASVASLTLTTTECAIAEKPEPKOMPFGGMMG 540
Db 481 ATGEYVDMVEAGVVDPTKVTTRIALQNAASVAGLLTTTDAVAEL--PKEDAPMPGMPGG 538

Qy 541 MGGMDGM 547
Db 539 MGGM-GM 544

```

## RESULT 13

```

Q62182 ID Q62182 PRELIMINARY; PRT; 550 AA.
AC Q62182;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Chaperonin, 60 kDa.
GN Name=groEL; ORFName=BMA2001;
OS Burkholderia mallei ATCC 23344.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=243160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23344;
RA Niernan W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T., Ulrich R.L., Ronning C.M., Brinkac L.M., Daugherty S.C.,
RA Davidson T.D., Deboy R.T., Dimitrov G., Dodson R.J., Durkin A.S.,
RA Gwinn M.L., Haft D.H., Khouri H., Kolonay J.F., Madupu R.,
RA Mohammed J., Nelson W.C., Radune D., Romero C.M., Sarria S.,
RA Selengut J., Shambhoo C., Sullivan S.A., White O., Yu Y., Zafar N.,
RA Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14247-14251(2004).
DR EMBL; CP000010; AAU50008.1; --
SQ SEQUENCE 550 AA; 57522 MW; E9E5028709EBE66E CRC64;

```

```

Query Match 69.0%; Score 1868.5; DB 2; Length 550;
Best Local Similarity 68.2%; Pred. No. 2.9e-80;
Matches 373; Conservative 75; Mismatches 96; Indels 3; Gaps 2;

```

```

Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVTLPKGRNVNIEKSPVITKDGVSVAKEI 60
Db 1 MAADKVVFGDSARAKVEGVNLANAVKVTLPKGRNVNIEKSPVITKDGVSVAKEI 60

Qy 61 ELEDKFNMGQMKVEKAPKTSIDAGDGTATVLAQIAVEGVKLVAAAGNPNMAIKRGI 120
Db 61 ELKDKLQNMGMQMKVEKASTSDNAGDGTATVLAQISVREGVKLVAAAGNPNMDLKRGI 120

Qy 121 DKAVAVVTKELSDITKTRDQKEIAQVGTISANSDDTTIGNIABAMAKVKGKGVITVEEA 180
Db 121 DKAVAAVEELKKISKCTTNKEIAQVGAISANSDDSSIGDRIAEAMDKVKGKGVITVEDG 180

```

```

Qy 181 KGLTTLDVVGKMPDRGYLSPYFVTNPKVWCVDLNPYILCNKKITSMKMDLPILEQV 240
Db 181 KSLADELDVVGKMPDRGYLSPYFVTNPKVQVAVLDNPFVLLHDKVSNIRDLPLVLEQV 240

Qy 241 AKVNRPLLIITAEVDVEGALATLVNKLKRGALQVAVKAPGGERKAWLEDIAITLTGGEA 300
Db 241 AKAGRPLLIITAEVDVEGALATLVNNGIRGLKTVAVKAPGGERKAWLEDIAITLTGGQV 300

Qy 301 IFEDGKLENVSLSSLTAKRVVIDKENTTIVDGAGKSEDIKARVKQIRAIQIBETSSDY 360
Db 301 IAEETGLTLEKATLAEGLQAKRIEVRKENTTIIDGAGEAASIEARVKQIRQIEEATSDY 360

Qy 361 DREKQERLAKLVGGVAVIHVGAATETEMKEDKDRVEDALNATRAAEEGIVPGGGTAFV 420
Db 361 DREKQERLAKLVGGVAVIKVGAATEVEMKEKARVEDALHATRAAEEGIVAGGGVALI 420

Qy 421 RSIKVLDDIKPADDELALGNIIRRSLEELPQIRQIAANAGVEGSIIVKVRPKDGFNFNA 480
Db 421 RARTAIAGLTGADQAGIKIVIRAMEEPLRQIVITNGGEEASVVVAAVAAGKNYGINA 480

Qy 481 ASGEYEDLIKAGVIDPKKVTTRIALQNAASVASLTLTTTECAIAEKPEPKOMPFGGMMG 540
Db 481 ATGEYVDMVEAGVVDPTKVTTRIALQNAASVAGLLTTTDAVAEL--PKEDAPMPGMPGG 538

Qy 541 MGGMDGM 547
Db 539 MGGM-GM 544

```

## RESULT 14

```

CH60_RHOPA ID CH60_RHOPA STANDARD; PRT; 546 AA.
AC Q93MH1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
GN Name=groL; Synonyms=groEL;
OS Rhodospseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodospseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RA Xiao M., Zhu C.R., Qian X.M., Zheng P., Chen Y.Y.;
RT "Cloning and sequencing of the groEL operon of Rhodospseudomonas palustris.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Prevents misfolding and promotes the refolding and proper assembly of unfolded polypeptides generated under stress conditions (By similarity).
CC -!- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of 7 subunits (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the chaperonin (HSP60) family.

```

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

```

```

DR EMBL; AF406639; AAK94943.1; --
DR HSP; Q92462; IIOK.
DR HAMAP; MF_00600; --; 1.
DR InterPro; IPR001844; Chaperin Cpn60.
DR InterPro; IPR002423; Cpn60/TCP-1.
DR InterPro; IPR008950; GroEL-ATPase.
DR Pfam; PF00118; Cpn60 TCP1; 1.
DR PRINTS; PR00298; CHAPERONIN60.
DR PRINTS; PR00304; TCOMPLEXTCP1.

```

```
DR PROSITE; PS00296; CHAPERONINS_CPN60; 1.
KW ATP-binding; Chaperone
SQ SEQUENCE 546 AA; 57937 MW; 163C7115D68FF8 CRC64;

Query Match 69.0%; Score 1868; DB 1; Length 546;
Best Local Similarity 68.2%; Pred. No. 3e-80;
Matches 374; Conservative 75; Mismatches 95; Indels 4; Gaps 2;

Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVLGPKGRNVVIEKSPGSPVITKDGVSVAKEI 60
Db 1 MAAKDVFEFDDARDMLRGVWIDKADAVKVLGPKGRNVVIEKSPGSPVITKDGVSVAKEI 60
Qy 61 ELEDKFNMGAAQVKEVAPKTSIDAGDGTATVLAQAIYREGVKLVAAAGNPMIAIKGI 120
Db 61 ELSKDFENMGAAQVKEVASTNDAGDGTATVLAQAIYREGVKLVAAAGNPMIAIKGI 120
Qy 121 DKAVVAVTKELSDITKPTDQKEIAQVGTISANSDDTTIGNIIAEMAKVGGVITVEEA 180
Db 121 DLATAKVVESIKAASRPVNDQHEVAQVGTISANGEAQIGRFIADAMQKVGNEGVIIVTEN 180
Qy 181 KGLETLDDVVEGKFDGRLSPYFVTNPEKVMCELDNPYILCNEKKITSMKMDLPILEQV 240
Db 181 KGLETEVEVVEGMDGRLSPYFVTNADKMTAELEDVFIILHKKLSLSIQPMVPLLESV 240
Qy 241 AKVNRPLLIITAEVDEGEALATLVVVKLRGALQVAVKAPGFGERRKAMLEDIAILLTGGEA 300
Db 241 IQAQRPLLIITAEVDEGEALATLVVVKLRGGLKIAAVKAPGFGDRKAMLEDIAILLTGQV 300
Qy 301 IFEDRGIKLENVSLSSIGTKARVVVDKENTTIIVDAGKSEDIKARVKQIRAEIETSSDY 360
Db 301 ISEDLGKLENVITDMLGRKVKVINKDNTTIIVDAGKAEIEARVSQIRTOIEETSSDY 360
Qy 361 DREKLOERLAKIVGGVAVIHVGAATEMEKEDKORVEDALNATRAAEEGIVPGGGTAFV 420
Db 361 DREKLOERLAKIVGGVAVIRVGGTMEVVEKEDKORVEDALNATRAAEEGIVVGGVALLI 420
Qy 421 RSIKVLDDIIPADDDDELALGNIIRRSLEELRQIAANAGYEGSVIVKVEPKDGFGEFN 479
Db 421 QAGKVLGLTCENPDQAGITIVRALEAPLQIAQNAQVDSVAVAGKVEDSDDKAFGFN 480
Qy 480 AASGEYEDLIKAGVIDPKKVTTRIALQNAASVASLLLTTECAIAEKPKPKDMPGGMG 539
Db 481 AQTEYGMFKFQVIDPAKVVRTALEDAASVASLLITTEAMIADKPSPKS---APAGGMG 537
Qy 540 GMGMDGM 547
Db 538 GMGMDGM 545

RESULT 15
CH60_BURTH
ID _CH60_BURTH STANDARD; PRT; 546 AA.
AC PS8723;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (groEL protein).
GN Name=groEL; Synonyms=groEL;
OS Burkholderia thailandensis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=57975;
RN [1]
RP SEQUENCE FROM N.A.
RA Woo P.C.Y., Woo G.K.S., Yuen K.Y.;
RT "Burkholderia thailandensis chaperonin GroEL gene, complete sequence."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Prevents misfolding and promotes the refolding and proper assembly of unfolded polypeptides generated under stress conditions (By similarity).
CC -!- SUBUNIT: Oligomer of 14 subunits composed of two stacked rings of 7 subunits (By similarity).
```

Search completed: April 12, 2005, 09:37:29  
Job time : 72 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 09:32:12 ; Search time 18 Seconds  
(without alignments)  
2929.266 Million cell updates/sec

Title: US-09-077-574A-2  
Perfect score: 2708  
Sequence: 1 MASKEILFDKAREKLSRGV.....KDMPPGGGGMGGMDGMY 548

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1866	68.9	551	2 JC2562	chaperonin groELx
2	1864	68.8	545	2 C95311	groEL2 chaperonin
3	1863	68.8	545	2 JN0509	heat shock protein
4	1862.5	68.8	547	2 F82783	60kDa chaperonin X
5	1852.5	68.4	546	2 B47073	chaperonin groEL -
6	1845	68.1	550	2 A41468	60K heat shock pro
7	1844.5	68.1	546	2 S65596	heat shock protein
8	1843.5	68.1	547	2 I40331	Cpn60 protein (Gro
9	1843.5	68.1	552	2 S39765	chaperonin 60 - Co
10	1841	68.0	542	2 F95967	probable heat choc
11	1840.5	68.0	544	2 B97442	60K chaperonin (pr
12	1840.5	68.0	544	2 AD2660	60 KDA chaperonin
13	1828.5	67.5	544	2 S23918	groEL protein - Ag
14	1825.5	67.4	547	2 B83098	groEL protein PA43
15	1823	67.3	546	2 S23347	groEL - Brucella a
16	1823	67.3	550	2 S35309	heat shock protein
17	1822	67.3	546	2 I40342	heat shock protein
18	1817.5	67.1	544	2 B36917	heat shock protein
19	1816.5	67.1	541	2 JN0512	heat shock protein
20	1814.5	67.0	544	2 H81964	chaperonin 60k su
21	1814.5	67.0	549	2 JC2564	heat shock protein
22	1814	67.0	546	2 AG3640	60K chaperonin gro
23	1812.5	66.9	545	2 S47530	chaperonin groEL -
24	1810.5	66.9	544	2 S37039	groEL protein - Ba
25	1808.5	66.8	544	2 C81031	chaperonin, 60 kDa
26	1805.5	66.7	547	2 B43606	heat shock protein
27	1802.5	66.6	544	2 S61300	heat shock protein
28	1799	66.4	546	2 B54539	heat shock protein
29	1794.5	66.3	544	2 S61302	heat shock protein

30	1794	66.2	546	2 S35311	heat shock cognate
31	1791.5	66.2	547	2 B87334	chaperonin, 60 kDa
32	1791	66.1	541	2 S61301	heat shock protein
33	1789.5	66.1	541	2 JN0511	heat shock protein
34	1788.5	66.0	544	2 S61303	heat shock protein
35	1786	66.0	545	2 C70489	groEL - Aquifex ae
36	1780.5	65.7	544	2 JC6063	chaperonin groEL -
37	1780.5	65.7	544	2 B83720	class I heat-shock
38	1774.5	65.5	545	2 S51563	heat shock protein
39	1773	65.5	544	1 B43827	chaperonin groEL -
40	1765	65.2	546	2 S34938	heat shock protein
41	1761.5	65.0	544	2 B82048	chaperonin, 60 kd
42	1761	65.0	548	2 AG0043	60 kDa chaperonin
43	1759	65.0	548	2 C64076	chaperonin groEL -
44	1755.5	64.8	539	2 B49855	heat shock protein
45	1755	64.8	545	2 G81328	60 kD chaperonin (

ALIGNMENTS

RESULT 1

JC2562

Chaperonin groELx protein - Amoeba proteus

C/Species: Amoeba proteus

C/Date: 17-May-1995 #sequence\_revision 14-Jul-1995 #text\_change 09-Jul-2004

C/Accession: JC2562

R/Abn, T.I.: Lim, S.T.; Leeu, H.K.; Lee, J.E.; Jeon, K.W.

Gene 148[128], 43-49, 1994

A/Title: A novel strong promoter of the groEx operon of symbiotic bacteria in Amoeba pro

A/Reference number: JC2561

A/Note: due to a typographical error the volume number 148 appears as 128

A/Accession: JC2562

A/Molecule type: DNA

A/Residues: 1-551 <AHN>

A/Cross-references: UNIPROT:P26004; GB:M86549; NID:g155400; PIDN:AAC09381.1; PID:g155402

C/Comment: This protein is involved in the assembly of oligomeric protein complexes, and

C/Genetics:

A/Gene: groELx

C/Superfamily: chaperonin groEL

C/Keywords: molecular chaperone

Query Match 68.9%; Score 1866; DB 2; Length 551;  
Best Local Similarity 67.9%; Pred. No. 3e-82;  
Matches 373; Conservative 74; Mismatches 96; Indels 6; Gaps 2;

Qy 3 SKELFDKAREKLSRGVDKLANAVKTLGPKGRNVVIEKSPVITKQGVSVAKIEL 62

Db 2 AKELRFQDDARQQMLAGVNLADRVKATMGPSGRNVVLSRFGAPTITKQGVSVAKIEIF 61

Qy 63 EDKFNMGAMVKEVAPKPTSDIAGDGTITATVLAQAITYREGVKLVAAAGRNPMATIKRGIDK 122

Db 62 ENRFKNMGAMVKEVAAKTSDTAGDGTITATVLAARSIVVEGHKAVAAGMNPMDLKRIGDK 121

Qy 123 AVAVTKELSDITYKPRDQKEIAQVGTISANSDDTTIGNIIAABAMAKVKGKGVITVEBAKG 182

Db 122 AVTAITKELQKMGKPKDGRKAIQAQVGTISANSQDAIGSIIAEAMEKVGKGVITVEDGNG 181

Qy 183 LETTLDVVEGKMDRGVLSYFVVTNPEKMCVDLNPVILCNEKKTSMKMDMLPILEOVAK 242

Db 182 LENELSVVEGMDRGVLSYFVVTNPEKMCVDLNPVILCNEKKTSMKMDMLPILEOVAK 241

Qy 243 VNRPLLIADVVEGEALATLVVVKLRGALQVAVKAPFGFGERRKAMLEADIAITGGBAIF 302

Db 242 SGRPLLIADVVEGEALATLVVNMGRGIVKCAVKAFFGDRRKAMLQDAILTNGQVIS 301

Qy 303 EDGIGKLENVSLSSLTAKRWIDKENTTIVDGAQKSEDIKARVKQIRAQIEETSSDYDR 362

Db 302 EBITGTSLETASLSLGTAKRWIDKENTTIVDGAQKSEDIKARVKQIRAQIEETSSDYDR 361

Qy 363 EKLQERLAKLVGGVAVIHVGAATETEMKEDKRVEDALNATRAAVERGIVPGGCTAFVRS 422

Db 362 EKLQERLAKLVGGVAVIHVGAATETEMKEDKRVEDALNATRAAVERGIVPGGCTAFVRS 421

```
QY 423 IKVLDDIKPADDDELAGNIIRSLRLEPLRQIRANAGVEGSIIVVEKYREPKDGFNAAS 482
Db 422 QKVLDGKGNADQDMGINIIRRAEPLRQIRVANAGYESSIVVKNVAEHKDNFGFNAAT 481
QY 483 GEYEDLIKAGVIDPKKYTRIALQNAASVASLLLTTECAIAEKPFPKDMMPG----GGM 538
Db 482 GOYDVMVEGILDPTKVRTALQNAASVRSILMLTTECWADL--PKKDEGAGAGDNGGM 539
QY 539 GGMGMDGM 547
Db 540 GGMGMDGM 548

RESULT 2
C95311
groEL2 chaperonin [imported] - Sinorhizobium meliloti (strain 1021) magaplasmid pSymA
C/Species: Sinorhizobium meliloti
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C/Accession: C95311
R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A/Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A/Reference number: A95262; MUID:21396509; PMID:11481432
A/Accession: C95311
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-545 <R>
A/Cross-references: UNIPROT:Q922Q4; GB:AB06469; PIDN:AAK65053.1; PID:g14523485; GSPDB:Q
A/Experimental source: strain 1021, magaplasmid pSymA
R/Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A/Reference number: A96039; MUID:21368234; PMID:11474104
A/Contents: annotation
C/Genetics:
A/Gene: groEL2
A/Genome: plasmid
C/Superfamily: chaperonin groEL

Query Match 68.8%; Score 1864; DB 2; Length 545;
Best Local Similarity 70.5%; Pred. No. 3.7e-82;
Matches 387; Conservative 59; Mismatches 97; Indels 6; Gaps 4;

QY 1 MASKEILFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSPGSPVITKDGVSVAKEI 60
Db 1 MAAKEVFGRSAREKMLRGVDILADAVKVTLPKGRNVVIDKSGAPRITKDGVTVAKEI 60
QY 61 ELEDKFNMGAAQWKEVAPKTSIDAGDGTATTATVLAQAIYREGVKLVAAAGNPMIAIKRGI 120
Db 61 ELEDKFNMGAAQWREVASKTNDIAGDGTATTATVLAQAIYREGAKAIVAAGNPNMDLKRG 120
QY 121 DKAVAVTKELSDITKTROKETAQVGTISANSDDTTIGNIIAEAMAKVGKGGVITVEEA 180
Db 121 DLVAEVVKOLLAKAKKINTSDEVAQVGTISANGEKQIGLDIAEAMQKVGNEGVITVEEA 180
QY 181 KGLTLDVVEGKMFDRGYLSPYFVTNPEKVCBLDNPYILCNKKTITSMKMDLPILAEQV 240
Db 181 KTAETELEVVEGMOFDRGYLSPYFVTNPEKXWADLEDAFILLHEKLSNLQMLPVLAEV 240
QY 241 AKVNRPLLIITAEDEVEGALATLVNKLRGALQVAVKAPGFGRRKAMLEDAIILTGEA 300
Db 241 VQTGKPLLIITAEDEVEGALATLVNKLRGGLKIAAIVKAPGFGDRRKAMLEDAIILTGT 300
QY 301 IFEDRGIKLENVSLISLGTAKRVVVDKENTTIVDGAGKSDIEDIKARVKQIRAOIETSSDY 360
Db 301 ISEDLGKLESVTLMDLGRACKVSIKENTTIVDGAGKSDIEGRVAQIKAOIETSSDY 360
QY 361 AKVNRPLLIITAEDEVEGALATLVNKLRGALQVAVKAPGFGRRKAMLEDAIILTGEA 420
Db 241 VQTGKPLLIITAEDEVEGALATLVNKLRGGLKIAAIVKAPGFGDRRKAMLEDAIILTGT 300
QY 301 IFEDRGIKLENVSLISLGTAKRVVVDKENTTIVDGAGKSDIEDIKARVKQIRAOIETSSDY 360
Db 301 ISEDLGKLESVTLMDLGRACKVSIKENTTIVDGAGKSDIEGRVAQIKAOIETSSDY 360
```

```
QY 361 DREKLQERLAKLVGGVAVIHVGAATETEMKEKKDRVEDALNATRAAAVEEIVPGGGTAFV 420
Db 361 DREKLQERLAKLVGGVAVIRVGATEVEVEKEDRIDDALNATRAAAVQEGIVPGGGVALL 420
QY 421 R-SIKVLDDIKPADDDELAGNIIRSLRLEPLRQIRANAGVEGSIIVVEKYREPK-DGFGF 478
Db 421 RSSVKI--TVKGENDDQDAGVIVRRALQSPARQIVENAGDEASIVVGKILEKNTDDFGY 478
QY 479 NAASGEVEDLIKAGVIDPKKYTRIALQNAASVASLLLTTECAIAEKEPEPKDMMPGGM 538
Db 479 NAQTGEYDGMIAWGIIDPVKVRTALQDAASVASLLITTEAMIABL--PKKDAAMPGGM 536
QY 539 GGMGMDGM 547
Db 537 GGMGMDGM 545

RESULT 3
JN0509
heat shock protein groEL (clone Rhz A) - Rhizobium meliloti
N/Alternate names: chaperonin groEL protein
C/Species: Rhizobium meliloti
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C/Accession: JN0509
R/Rusanganwa, E.; Gupta, R.S.
Gene 126, 67-75, 1993
A/Title: Cloning and characterization of multiple groEL chaperonin-encoding genes in Rhiz
A/Reference number: JN0509; MUID:93231539; PMID:8097179
A/Accession: JN0509
A/Molecule type: DNA
A/Residues: 1-545 <R>
A/Cross-references: UNIPROT:P35469; GB:M94192; NID:g1522233; PIDN:AAA26285.1; PID:g152233
C/Comment: This protein plays a role in protein folding and in the extracellular transpo
C/Genetics:
A/Gene: groEL
C/Superfamily: chaperonin groEL
C/Keywords: heat shock; molecular chaperone; stress-induced protein

Query Match 68.8%; Score 1863; DB 2; Length 545;
Best Local Similarity 70.5%; Pred. No. 4.1e-82;
Matches 387; Conservative 58; Mismatches 98; Indels 6; Gaps 4;

QY 1 MASKEILFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSPGSPVITKDGVSVAKEI 60
Db 1 MAAKEVFGRSAREKMLRGVDILADAVKVTLPKGRNVVIDKSGAPRITKDGVSVAKEI 60
QY 61 ELEDKFNMGAAQWKEVAPKTSIDAGDGTATTATVLAQAIYREGVKLVAAAGNPMIAIKRGI 120
Db 61 ELEDKFNMGAAQWREVASKTNDIAGDGTATTATVLAQAIYREGAKAIVAAGNPNMDLKRG 120
QY 121 DKAVAVTKELSDITKTROKETAQVGTISANSDDTTIGNIIAEAMAKVGKGGVITVEEA 180
Db 121 DLVAEVVKOLLAKAKKINTSDEVAQVGTISANGEKQIGLDIAEAMQKVGNEGVITVEEA 180
QY 181 KGLTLDVVEGKMFDRGYLSPYFVTNPEKVCBLDNPYILCNKKTITSMKMDLPILAEQV 240
Db 181 KTAETELEVVEGMOFDRGYLSPYFVTNPEKXWADLEDAFILLHEKLSNLQMLPVLAEV 240
QY 241 AKVNRPLLIITAEDEVEGALATLVNKLRGALQVAVKAPGFGRRKAMLEDAIILTGEA 300
Db 241 VQTGKPLLIITAEDEVEGALATLVNKLRGGLKIAAIVKAPGFGDRRKAMLEDAIILTGT 300
QY 301 IFEDRGIKLENVSLISLGTAKRVVVDKENTTIVDGAGKSDIEDIKARVKQIRAOIETSSDY 360
Db 301 ISEDLGKLESVTLMDLGRACKVSIKENTTIVDGAGKSDIEGRVAQIKAOIETSSDY 360
QY 361 DREKLQERLAKLVGGVAVIHVGAATETEMKEKKDRVEDALNATRAAAVEEIVPGGGTAFV 420
Db 361 DREKLQERLAKLVGGVAVIRVGATEVEVEKEDRIDDALNATRAAAVQEGIVPGGGVALL 420
QY 421 R-SIKVLDDIKPADDDELAGNIIRSLRLEPLRQIRANAGVEGSIIVVEKYREPK-DGFGF 478
Db 421 RSSVKI--TVKGENDDQDAGVIVRRALQSPARQIVENAGDEASIVVGKILEKNTDDFGY 478
```







I40331  
Cpn60 protein (GroEL) - Bordetella pertussis  
C;Species: Bordetella pertussis  
C;Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 09-Jul-2004  
C;Accession: I40331  
R;Fernandez, R.C.; Weiss, A.A.  
Gene 159, 151-152, 1995  
A;Title: Cloning and sequencing of the Bordetella pertussis cpn10/cpn60 (groESL) homolog  
A;Reference number: I40330; MUID:95309719; PMID:7789805  
A;Accession: I40331  
A;Molecule type: DNA  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Residues: 1-547 <RES>  
A;Cross-references: UNIPROT:P48210; EMBL:U12277; NID:G968918; PIDN:AAA74967.1; PID:G968918  
C;Genetics:  
C;Gene: cpn60  
C;Superfamily: chaperonin groEL

Query Match 68.1%; Score 1843.5; DB 2; Length 547;  
Best Local Similarity 67.4%; Pred. No. 3.6e-81;  
Matches 370; Conservative 74; Mismatches 98; Indels 7; Gaps 2;

Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSPGSPVITKDGVSVAKEI 60  
Db 1 MAAKQVLFADEARVIRVGVNLANAVKVTLPKGRNVVILERSFGAPTITKDGVSVAKEI 60

Qy 61 ELEDKFNMGAGQVMKVEAPKTSIDAGDGTATTATVLAQAIYREGVKLVAAAGNPNMAIKRGI 120  
Db 61 ELKDKFENIGAQLVKDVASKTSIDNAGDGTATTATVLAQAVQEGKLVYAAAGNPNIDLRGI 120

Qy 121 DKAVVAVTKELSDITKTRDROKETAQVGTISANSDDTTIGNIIAAMAKVKGKGVITVEEA 180  
Db 121 DKAVAAVVEELKLSKPVTTSKETAQVGSISANSDSIGQIIADAMDVKVKEGVTIVEDG 180

Qy 181 KGLETTLDVVEGKMFDRGYLSPYFVTNPEKVMCELDNPYLNCNEKKTITSMKMLPILQOV 240  
Db 181 KSLNELDVVEGQMFDRGYLSPYFINSPEKQVQALDDPYVLIYDKKVSINRDLPLVLEQV 240

Qy 241 AKVNRPLIIIAEDVEGEALATLVVNNKLRGALQVAVKAPGFGERRKAMLEDIAITLGEA 300  
Db 241 AKSSRPLIIIAEDVEGEALATLVVNNKIRGTLTAVKAPGFGDRRRKAMLEDIAITLGGTV 300

Qy 301 IFEDRGKLENVSLSSIGTAKRVVVDKENTTIIDGAGKSEDIKARVKQIRAIQIBETSDDY 360  
Db 301 ISEETGMSLEKATLDLQKARIKAVAKENTTIIDGAGDGKSEIARVKQIRAIQIBETSDDY 360

Qy \*361 DREKQLERLAKLVGGVAVIHVGAATETEMKEKKDRVEDALNATRAAAVEEGIVPGGGTAFV 420  
Db 361 DREKQLERVAKLGGVAVIRVGAATEVEMKEKKARVEDALHATRAAAVEEGVPPGGVALL 420

Qy 421 RSIKVLDDIKPADDELGLAGNIIIRRSLEELPQIAANAGYEGSIVVEKVPKDPGFGFNA 480  
Db 421 RAKQAITGLKGDATADQAGIKLIIRAVEEPLRTIVTNAGDEASVVVNTVLNGKNGYNA 480

Qy 481 ASGEYEDLIKAGVIDPKVTRIALQNAASVASILLITTECAIAE----KPEPKDMPNPGG 536  
Db 481 ATGEYGLVQGVLDPTKVTRTALQNAASVASLLLTAAAVVELMENKPAAPAMP---G 537

Qy 537 GMGGMGGMD 545  
Db 538 GMGGMGGMD 546

RESULT 9  
S39765  
chaperonin 60 - Coxiella burnetii  
C;Species: Coxiella burnetii  
C;Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C;Accession: S39765  
R;Vodkin, M.H.; Williams, J.C.  
J. Bacteriol. 170, 1227-1234, 1988  
A;Title: A heat shock operon in Coxiella burnetii produces a major antigen homologous to  
A;Reference number: S39764; MUID:88139182; PMID:3343219

A;Accession: S39765  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-552 <VD>  
A;Cross-references: UNIPROT:P19421; GB:M20482; NID:G144996; PIDN:AAA23309.1; PID:G144998  
C;Superfamily: chaperonin groEL

Query Match 68.1%; Score 1843.5; DB 2; Length 552;  
Best Local Similarity 66.5%; Pred. No. 3.6e-81;  
Matches 365; Conservative 82; Mismatches 99; Indels 3; Gaps 3;

Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSPGSPVITKDGVSVAKEI 60  
Db 1 MAAKVLFSEHVLHMSRGEVLANAVKVTLPKGRNVVLDKSGFAPITTKDGVSVAKEI 60

Qy 61 ELEDKFNMGAGQVMKVEAPKTSIDAGDGTATTATVLAQAIYREGVKLVAAAGNPNMAIKRGI 120  
Db 61 ELEDKFNMGAGQVMKVEASRTSDDAGDGTATTATVLAQAILVEGKAVIAGNPNMDLKRGI 120

Qy 121 DKAVVAVTKELSDITKTRDROKETAQVGTISANSDDTTIGNIIAAMAKVKGKGVITVEEA 180  
Db 121 DKAVTAAVSELKTKSKPKDQKAIQVGTISANSDDKISIGDIIAEMEKVKGKGVITVBDG 180

Qy 181 KGLETTLDVVEGKMFDRGYLSPYFVTNPEKVMCELDNPYLNCNEKKTITSMKMLPILQOV 240  
Db 181 SGLNELLEVVEGQMFDRGYLSPYFINNQNMSELENPFILLVDKKISINRELIPLENV 240

Qy 241 AKVNRPLIIIAEDVEGEALATLVVNNKLRGALQVAVKAPGFGERRKAMLEDIAITLGEA 300  
Db 241 AKSRPLIIIAEDVEGEALATLVVNNIRGVVAAVKAAPGFGDRRKAMLDIAVLTGGKV 300

Qy 301 IFEDRGKLENVSLSSIGTAKRVVVDKENTTIIDGAGKSEDIKARVKQIRAIQIBETSDDY 360  
Db 301 ISEVGLSLAASLDDIGSARVVVTKDDTTIIDGSGDAGDIKRVQIRKEIENSDDY 360

Qy 361 DREKQLERLAKLVGGVAVIHVGAATETEMKEKKDRVEDALNATRAAAVEEGIVPGGGTAFV 420  
Db 361 DREKQLERLAKLVGGVAVIKVGAATEVEMKEKKARVEDALHATRAAAVEEGVPPGGVALL 420

Qy 421 RSIKVLDDIKPADDELGLAGNIIIRRSLEELPQIAANAGYEGSIVVEKVPKDPGFGFN 479  
Db 421 RVLSLDSVEVENEDQVGEIARRAMAYPLSQIVKTVQVAAVADKVLNKHNDVNGYN 480

Qy 480 ASGEYEDLIKAGVIDPKVTRIALQNAASVASLLITTECAIAEAKPKDMPNPGG-M 538  
Db 481 AATGEYDGMTIEMGLDPTKVTRTALQNAASVAGLITTECVMTAPK-KKEESPPGGGDM 539

Qy 539 GMGGMGGMD 547  
Db 540 GMGGMGGMD 548

RESULT 10  
P59567  
probable heat shock protein groEL [imported] - Sinorhizobium meliloti (strain 1021) mag  
C;Species: Sinorhizobium meliloti  
C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: P59567  
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo  
A;Reference number: A95842; MUID:21396508; PMID:11481431  
A;Accession: P59567  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-542 <KUR>  
A;Cross-references: UNIPROT:P35471; GB:AL591985; PIDN:CAC49406.1; PID:G15140892; GSPDB:G  
A;Experimental source: strain 1021, megaplasmid pSymB  
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.K.; Yeh, K.  
A;Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
A;Reference number: A96039; MUID:21368234; PMID:11474104

C;Contents: annotation  
C;Genetics:  
A;Gene: groEL5; Smb21566  
A;Genome: plasmid  
C;Superfamily: chaperonin groEL

Query Match 68.0%; Score 1841; DB 2; Length 542;  
Best Local Similarity 68.3%; Pred. No. 4.6e-81;  
Matches 373; Conservative 71; Mismatches 96; Indels 6; Gaps 2;

```
QY 1 MASKEILFDKAREKLSRGVVDKLANAVKVTLPKGRNVVIEKSPGSPVITKDGYSVAKEI 60
DB 1 MAKEVFGQDARMLRGVVDKLANAVKVTLPKGRNVVIEKSPGSPVITKDGYSVAKEI 60
QY 61 ELEDKFNMGQAQVKEVAPKTSIAGDGTATVLAQAIYREGVKLVAAAGNPMIAIKRGI 120
DB 61 ELEDKFNMGQAQVKEVAPKTSIAGDGTATVLAQAIYREGVKLVAAAGNPMIAIKRGI 120
QY 121 DKAVAVTKELSDITKPTROKELIAQVGTISANSDDTTIGNIIAEMAKVKGKGVITVEEA 180
DB 121 DLAVDAVVKELKNNARKISKNSIAQVGTISANSDDTTIGNIIAEMAKVKGKGVITVEEA 180
QY 181 KGLETTLDVVEGKFDGYSLSPYFTNPKVWCELDNPYILCNEKKITSMKMDLPILQEV 240
DB 181 KTAETELEVVEGQFDGYSLSPYFTNPKVWCELDNPYILCNEKKITSMKMDLPILQEV 240
QY 241 AKVNRPLIIIAEDVEGEGALATLVNKLRLGALQVAVKAPGFGERRKAMLEDAIILTGTV 300
DB 241 VQSGKPLIIIAEDVEGEGALATLVNKLRLGALQVAVKAPGFGERRKAMLEDAIILTGTV 300
QY 301 IFEDRGKLENVSLSSIGTAKRVVVDKENTTIIDGAGSKSDIKARVKQIRAQIBETSSDY 360
DB 301 VSEDGLKLESVTLMDLGRKAKVSIKENTTIIDGAGSKSDIKARVKQIRAQIBETSSDY 360
QY 361 DREKQLERLAKLVGGVAVIHVGAATEMEKEDKDRVEDALNATRAAEEGIVPGGGTAFV 420
DB 361 DREKQLERLAKLVGGVAVIHVGGSTEVEKEKDRVEDALNATRAAEEGIVPGGGTAFV 420
QY 421 RSIKVLDDIPADDELALGNIIRRSLEPLRQIAANAGYEGSVIVBKVRPKD-GRGFN 479
DB 421 RAVKALDGLKTANNDRQVGVLDVRAIEAPVRQIAENAGAEGSVIVBKVRPKD-GRGFN 480
QY 480 AASGEYEDLLKAGVIDPKVTRIALQNAASVASILLTTECAIAEKPKPKDMPMPGGMG 539
DB 481 AQTNEYGDLVAMGVIDPAKVVRTALQDAASVAGLLVTTEAMIAEKPKKEAAPALPAG--- 537
QY 540 GMGGMD 545
DB 538 --GGMD 541
```

RESULT 11  
B97442  
60K chaperonin (protein cpn60) (groEL protein) [imported] - Agrobacterium tumefaciens (s  
C;Species: Agrobacterium tumefaciens  
C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
C;Accession: B97442  
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A;Reference number: A97359; MUID:21608551; PMID:11743194  
A;Accession: B97442  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-544 <KUR>  
A;Cross-references: UNIPROT:P30779; GB:AE007869; PIDN:AAK86491.1; PID:gl5155641; GSPDB:C  
C;Genetics:  
A;Gene: AGR\_C\_1220  
A;Map position: circular chromosome

C;Superfamily: chaperonin groEL

Query Match 68.0%; Score 1840.5; DB 2; Length 544;  
Best Local Similarity 69.2%; Pred. No. 4.9e-81;  
Matches 376; Conservative 69; Mismatches 95; Indels 3; Gaps 3;

```
QY 1 MASKEILFDKAREKLSRGVVDKLANAVKVTLPKGRNVVIEKSPGSPVITKDGYSVAKEI 60
DB 1 MAKEVFGRSAREKMLKGVDDIADAVKVTLPKGRNVVIEKSPGSPVITKDGYSVAKEI 60
QY 61 ELEDKFNMGQAQVKEVAPKTSIAGDGTATVLAQAIYREGVKLVAAAGNPMIAIKRGI 120
DB 61 ELEDKFNMGQAQVKEVAPKTSIAGDGTATVLAQAIYREGVKLVAAAGNPMIAIKRGI 120
QY 121 DKAVAVTKELSDITKPTROKELIAQVGTISANSDDTTIGNIIAEMAKVKGKGVITVEEA 180
DB 121 DLAVEVVDLQAKAKKINTSEEAQVGTISANGEROIGLDIAEAMQVGVNEGVIITVEEA 180
QY 181 KGLETTLDVVEGKFDGYSLSPYFTNPKVWCELDNPYILCNEKKITSMKMDLPILQEV 240
DB 181 KTAETELEVVEGQFDGYSLSPYFTNPKVWCELDNPYILCNEKKITSMKMDLPILQEV 240
QY 241 AKVNRPLIIIAEDVEGEGALATLVNKLRLGALQVAVKAPGFGERRKAMLEDAIILTGTV 300
DB 241 VQSGKPLIIIAEDVEGEGALATLVNKLRLGALQVAVKAPGFGERRKAMLEDAIILTGTV 300
QY 301 IFEDRGKLENVSLSSIGTAKRVVVDKENTTIIDGAGSKSDIKARVKQIRAQIBETSSDY 360
DB 301 ISEDLGKLESVTLMDLGRKAKVSIKENTTIIDGAGSKSDIKARVKQIRAQIBETSSDY 360
QY 361 DREKQLERLAKLVGGVAVIHVGAATEMEKEDKDRVEDALNATRAAEEGIVPGGGTAFV 420
DB 361 DREKQLERLAKLVGGVAVIHVGGSTEVEKEKDRVEDALNATRAAEEGIVPGGGTAFV 420
QY 421 RSIKVLDDIPADDELALGNIIRRSLEPLRQIAANAGYEGSVIVBKVRPKD-GRGFN 479
DB 421 RSSTKI--TVKGVNDQDQAGINIVRKALQSLVRQIAENAGDEASIVVGKILDKNEDNYGN 479
QY 480 AASGEYEDLLKAGVIDPKVTRIALQNAASVASILLTTECAIAEKPKPKDMPMPGGMG 538
DB 480 AQTGEYDGLIALGIVDPKVVRTALQNAASVASILLTTEAMIAELPKKESAMPQPGMG 539
QY 539 GGM 541
DB 540 GGM 542
```

RESULT 12

AD2660

60 KDA chaperonin [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004

C;Accession: AD2660

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan,

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, J.

ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AD2660

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-544 <KUR>

A;Cross-references: UNIPROT:P30779; GB:AE008688; PIDN:AAL41698.1; PID:gl7739044; GSPDB:G

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: groEL

A;Map position: circular chromosome

C;Superfamily: chaperonin groEL

Query Match 68.0%; Score 1840.5; DB 2; Length 544;

Best Local Similarity 69.2%; Pred. No. 4.9e-81;		
Matches 376; Conservative 69; Mismatches 95; Indels 3; Gaps 3;		
Qy	1 MASKEILLFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSPGSPVITKDGVSVAKEI 60	
Db	1 MAAEVKFGSAREKMLKGVLDLADAVKVTLPKGRNVVIEKSPGAPRITKDGVSVAKEI 60	
Qy	61 ELEDKPFNMGAQMVKVAPKTSITAGDGTATTATVLAQAIYREGVKLVAAGRNPMAIKRGI 120	
Db	61 ELEDKPFNMGAQLVREVASKNTDITAGDGTATTATVLAQAIYREGAKAVAAAGNPMDLKRG 120	
Qy	121 DKAVAVTKELSDITKPTROKETAQVGTISANSDDTIGNIIAEAMAKVGKGVITVEEA 180	
Db	121 DLAAEVVKDLQAKAKKINTSEEAQVGTISANGEROGLDIAEAMQVNEGVIITVEEA 180	
Qy	181 KGLFTTLDVVEGMPKFDRGYLSPYFVTNPKNVCELDNPPYILCNEKKITSMKMLPILEQV 240	
Db	181 KTAETELEVVEGMPKFDRGYLSPYFVTNPKNVADLEDAIILLHEKKLSNLQAMLPVLEAV 240	
Qy	241 AKVNRPLIIIAEDVEGALATLVNKLKGALOVVAVKAPGFGERRKAMLEDDIALLTGGEA 300	
Db	241 VQTCKPLIVIAEDVEGALATLVNKLKGLKIAAVKAPGFGDRRKAMLEDDIALLTGTV 300	
Qy	301 IFDRGKIKLENVSLSSLTAKRVVIDKENTTIVDGAQKSDIKARVKQIRAOIETEETSSDY 360	
Db	301 ISEDLGIKLESVTLMDLCKSKVISIKENTTIVDGAQKSDIEGRVVAQIRAOIETEETSDY 360	
Qy	361 DREKQLQRLAKLVGGVAVIHVGAAATEMKEKKORVEDALNATRAAAVEEGIVPGGTAFFV 420	
Db	361 DREKQLQRLAKLVGGVAVIRVGGSTEVEVKEKKRIDDALNATRAAAVQEGIVPGGVAL 420	
Qy	421 RSIKVLDDIKPADDDDELAGLINIRRSLEELPRLQIAANAGYEGSIIVVEKVRP-KDQFGFN 479	
Db	421 RSSTKI-TVKGVNDQDQAGINIRKALQSLVRQIAENAGDEASIVVGGKILDKKNDNYGN 479	
Qy	480 AASGEYEDLTAKAGVIDPKKVTRIALQNAASVASLLLTTECAIAEKPKPKDMP-MPGGGM 538	
Db	480 AQTGEYGDLTALGIVDPKVVRTALQNAASVASLLIITEAMIAELPKKESAMPQPGGM 539	
Qy	539 GGM 541	
Db	540 GGM 542	
RESULT 13		
S23918		
groEL protein - Agrobacterium tumefaciens		
C:Species: Agrobacterium tumefaciens		
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004		
C:Accession: S23918		
R:Segal, G.; Ron, E.Z.		
submitted to the EMBL Data Library, August 1992		
A:Description: Cloning and sequencing of the GroE operon of Agrobacterium tumefaciens C5		
A:Reference number: S23917		
A:Accession: S23918		
A:Status: preliminary		
A:Molecule type: DNA		
A:Residues: 1-544 <SSG>		
A:Cross-references: UNIPROT:P30779; EMBL:X68263; NID:g1019913; PIDN:CAA48331.1; PID:g390		
C:Superfamily: chaperonin groEL		
Query Match 67.5%; Score 1828.5; DB 2; Length 544;		
Best Local Similarity 68.9%; Pred. No. 1.9e-80;		
Matches 374; Conservative 70; Mismatches 96; Indels 3; Gaps 3;		
Qy	1 MASKEILLFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSPGSPVITKDGVSVAKEI 60	
Db	1 MAAEVKFGSAREKMLKGVLDLADAVKVTLPKGRNVVIEKSPGAPRITKDGVSVAKEI 60	
Qy	61 ELEDKPFNMGAQMVKVAPKTSITAGDGTATTATVLAQAIYREGVKLVAAGRNPMAIKRGI 120	
Db	61 ELEDKPFNMGAQLVREVASKNTDITAGDGTATTATVLAQAIYREGAKAVAAAGNPMDLKRG 120	

```
Db 181 SGLNELSVVEGMQFDRGYLSPYFVNKPDTHMAELDSPLLLLVDDKKINSIREMLPVLEAV 240
Qy 241 AKVNRPLLIITAEVDVEGALATLVNKLRGALQVAVKAPGFGERRKAMLEDIAIILTGEA 300
Db 241 AKAGRPLLIIVAEVDVEGALATLVNNGRGIVKVAARVAPGFGDRKRKAMLDIAIILTGTG 300
Qy 301 IFEDRGIKLENVSLSSIGTAKRVVIDKENTTIVDGAGKSEDIKARVKQIRAOIETSSDY 360
Db 301 ISEVGSLLEGATUEHGNKRVVINKENTTIIDAGVQADIEARVQIRKQIEETSSDY 360
Qy 361 DREKQLERLAKLVGGVAVIHVGAATETEMKEKDRVEDALNATRAAVEEGIVPGGGTAFV 420
Db 361 DREKQLERLAKLVGGVAVIKVGAATEVEMKEKARVEDALHATRAAVEEGVVPGGVALV 420
Qy 421 RSIKVLDDIKPADDELAGLNIIRSRLEBPLRQIAANAGYEGSIVVEKVRBPKDGFQFNA 480
Db 421 RALQAIIEGLKGDNEEQNVGIALLRVAVESPLRQIVANAGDEPSVVVDKVKQSGNGYFNA 480
Qy 481 ASGEYEDLIKAGVIDPKKVTTRIALONAAVASLILTTTECAIAEKPEPKDMPPGG--GM 538
Db 481 ATGVYGDNIENGILDDPAKVTIRSAQAASISGGLMITTEAMVAEIVE---DKPAMGGMPDM 537
Qy 539 GGMGGMGMDGM 547
Db 538 GGMGGMGGM 546

RESULT 15
S22347
groEL - Brucella abortus
C:Species: Brucella abortus
C>Date: 22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C:Accession: S22347
R:Gor, D.; Mayfield, J.E.
A:Title: Cloning and nucleotide sequence of the Brucella abortus groE operon.
A:Reference number: S22346; MUID:92182006; PMID:1347461
A:Accession: S22347
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-546 <GOR>
A:Cross-references: UNIPROT:P25967; EMBL:M82975; NID:g144109; PIDN:AAA22997.1; PID:g1441
C:Superfamily: chaperonin groEL

Query Match 67.3%; Score 1823; DB 2; Length 546;
Best Local Similarity 68.9%; Pred. No. 3.4e-80;
Matches 378; Conservative 64; Mismatches 99; Indels 8; Gaps 5;

Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVTILGPKGRNVVIEKSPGVITKDGVSVAKEI 60
Db 1 MAADVKFGRtareKMLRGVDILADAVKVTILGPKGRNVVIEKSPGAPRITKDGVSVAKEV 60
Qy 61 ELEDKFNMGAGQVMEVAPKTSIDTAGDGTATTATVLAQAIYREGVKLVAAGRNPMAIKRG 120
Db 61 ELEDKFNMGAGQVLEAVSKTNDTAGDGTATTATVLAQAIYQEGAKAVAAGNPMDLKRG 120
Qy 121 DKAVVATKELSDITKTROQKEIAQVGTISANSDDTTIGNIIAEAMAKVKGCVITVEEA 180
Db 121 DLAVNEVVAELLKAKKINTSEEVAQVGTISANAEEIGKMIAEAMQKMGVEGVTVEEA 180
Qy 181 KGLETTLDVVEGMKFDRLGYSLPYFVNTPEKMWCELDNPNYILCNEKKTITSMKMDLPILQV 240
Db 181 KTAETELEVVEGMQFDRGYLSPYFVTNPEKMWADLEDAYILLHEKLSNQLLPVLEAV 240
Qy 241 AKVNRPLLIITAEVDVEGALATLVNKLRGALQVAVKAPGFGERRKAMLEDIAIILTGEA 300
Db 241 VQTSKPLLIITAEVDVEGALATLVNKLRGGLKIAAVKAPGFGDRKRKAMLEDIAIILTGQV 300
Qy 301 IFEDRGIKLENVSLSSIGTAKRVVIDKENTTIVDGAGKSEDIKARVKQIRAOIETSSDY 360
Db 301 ISEDLGIKLSVTLDMUGRAKVSISKENTTIVDGAGKAEIDARVGQIKQIEETSSDY 360
Qy 361 DREKQLERLAKLVGGVAVIHVGAATETEMKEKDRVEDALNATRAAVEEGIVPGGGTAFV 420
```

```
Db 361 DREKQLERLAKLVGGVAVIRVGGAATEVEVEKDRVDDALNATRAAVEEGIVAGGGTALL 420
Qy 421 R-SIKVLDDIKPADDDDELAGLNIIRSRLEBPLRQIAANAGYEGSIVVEKVRB-PKDGFGF 478
Db 421 RASTKI--TAKGVNADQEGAGINIVRAIQAPARQITTNAGEEASVIVGKILENTSETFGY 478
Qy 479 NAASGEYEDLIKAGVIDPKKVTTRIALONAAVASLILTTTECAIAEKPEPKDMPPG--G 536
Db 479 NTANGEYGDLLISLGIIVDPVKVVRTALONAAVASAGLLITTEAMIAEL--PKKDAAPAGMPG 536
Qy 537 GGMGGMGMD 545
Db 537 GGMGGMGMD 545
```

Search completed: April 12, 2005, 09:39:16  
Job time : 21 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 09:38:59 ; Search time 136 Seconds  
(without alignments)

1337.754 Million cell updates/sec

Title: US-09-077-574A-2

Perfect score: 2708

Sequence: 1 MASKEILFDKAREKLSRGV.....KDMPPGGGGMGGMGMDGMY 548

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 331997259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1868.5	69.0	550	US-10-282-122A-50017	Sequence 50017, A
2	1862.5	68.8	547	US-10-369-493-17524	Sequence 17524, A
3	1857	68.6	540	US-10-282-122A-47926	Sequence 47926, A
4	1857	68.6	547	US-10-369-493-10014	Sequence 10014, A
5	1850.5	68.3	546	US-10-282-122A-49932	Sequence 49932, A
6	1843.5	68.1	547	US-10-282-122A-50796	Sequence 50796, A
7	1840.5	68.0	544	US-10-369-493-11760	Sequence 11760, A
8	1839.5	67.9	547	US-10-369-493-7917	Sequence 7917, Ap
9	1837.5	67.9	546	US-10-282-122A-68274	Sequence 68274, A
10	1835.5	67.8	542	US-10-369-493-10295	Sequence 10295, A
11	1835.5	67.7	541	US-10-369-493-9287	Sequence 9287, Ap
12	1833	67.7	544	US-10-369-493-15468	Sequence 15468, A
13	1833	67.7	544	US-10-369-493-16217	Sequence 16217, A

14	1832	67.7	545	15	US-10-282-122A-63193	Sequence 63193, A
15	1828.5	67.5	543	15	US-10-369-493-15265	Sequence 15265, A
16	1826	67.4	542	15	US-10-369-493-11922	Sequence 11922, A
17	1825.5	67.4	547	15	US-10-282-122A-66642	Sequence 66642, A
18	1820	67.2	545	16	US-10-375-010-28	Sequence 28, Appl
19	1814.5	67.0	544	15	US-10-282-122A-65715	Sequence 65715, A
20	1812.5	66.9	551	15	US-10-369-493-12330	Sequence 12330, A
21	1811	66.9	540	15	US-10-369-493-15838	Sequence 15838, A
22	1810	66.8	552	15	US-10-369-493-12170	Sequence 12170, A
23	1805.5	66.7	544	15	US-10-282-122A-65021	Sequence 65021, A
24	1796.5	66.3	549	15	US-10-282-122A-69386	Sequence 69386, A
25	1786	66.0	545	15	US-10-369-493-167	Sequence 167, App
26	1784.5	65.9	545	15	US-10-282-122A-48944	Sequence 48944, A
27	1780.5	65.7	544	15	US-10-369-493-17132	Sequence 17132, A
28	1778.5	65.7	545	14	US-10-228-167A-2	Sequence 2, Appli
29	1771.5	65.4	543	15	US-10-369-493-11916	Sequence 11916, A
30	1771	65.4	544	15	US-10-282-122A-46245	Sequence 46245, A
31	1770	65.4	536	15	US-10-369-493-4511	Sequence 4511, Ap
32	1770	65.4	540	15	US-10-369-493-7269	Sequence 7269, Ap
33	1768.5	65.3	544	15	US-10-282-122A-44818	Sequence 44818, A
34	1767.5	65.3	539	16	US-10-375-010-32	Sequence 32, Appl
35	1761.5	65.0	544	15	US-10-282-122A-77547	Sequence 77547, A
36	1761	65.0	548	15	US-10-282-122A-78505	Sequence 78505, A
37	1759	65.0	548	15	US-10-282-122A-58212	Sequence 58212, A
38	1755	64.8	545	15	US-10-282-122A-54546	Sequence 54546, A
39	1752	64.7	548	15	US-10-282-122A-75259	Sequence 75259, A
40	1752	64.7	548	15	US-10-282-122A-76037	Sequence 76037, A
41	1746	64.5	545	14	US-10-192-419-2	Sequence 2, Appli
42	1746	64.5	547	15	US-10-282-122A-67186	Sequence 67186, A
43	1746	64.5	548	10	US-09-415-849-1	Sequence 1, Appli
44	1746	64.5	548	10	US-09-276-455-10	Sequence 10, Appl
45	1746	64.5	548	15	US-10-282-122A-42859	Sequence 42859, A

#### ALIGNMENTS

#### RESULT 1

US-10-282-122A-50017  
; Sequence 50017, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: EITRA.034A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR FILING DATE: 2000-11-27





FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 47926  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Burkholderia cepacia  
US-10-282-122A-47926

Query Match 68.6%; Score 1857; DB 15; Length 540;  
Best Local Similarity 68.1%; Pred. No. 3.2e-119;  
Matches 369; Conservative 75; Mismatches 96; Indels 2; Gaps 1;  
QY 3 SKEILFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSFSGSPVITKDGVSVAKEI 62  
DB 1 AKDVFGDSARKNVGEVILANAVKVTLPKGRNVVILERSFGPTVTTKDGVSVAKEI 60  
QY 63 EDKFNENGAQWKEVAPKTSIADGDTTATVLAQAIYREGVLAAGRNPMATKRGIDK 122  
DB 61 KDLQNNGAQWKEVASKTSNADGDTTATVLAQSIYREGKVVASGMPMDLKRIGDK 120  
QY 123 AVAVATKELSDITKPTDQKEIAQVGTISANSDDTTIGNIIAEMAKVKGCVITVEBAK 182  
DB 121 AVAAAEVBLKISKPTTKEIAQVGSISANSDDSSIGDRIAEAMDVKVKEGVIIVEDGKS 180  
QY 183 LETTLDVVEGKMFDRGYLSPYFTVNPENKVCLENDNPYILCNEKKITSMKMDLPILQVAK 242  
DB 181 LADELVDVEGQMFDRGYLSPYFINNPKQVAVLNDPFLVLDKVKVSNIRDLPLVLEQVAK 240  
QY 243 VNRPLLIIEADVEGEALATLVNKLRGALQVAVKAPGFGERRKAMLEDAITLTGGVAIF 302  
DB 241 AGRPLLIIEADVEGEALATLVNIRGILKTVAVKAPGFGDRRKAMLEDAITLTGGVIA 300  
QY 303 EDGKILNVSLSLGTAKRVVIDKENTTIVDAGKSEDIKARVKQIRAOIETSSDYR 362  
DB 301 EETGLTEKATLAEKGAKRIEVCENTTIDGAGEAASIEARVKVQRAQIEEATS DYR 360  
QY 363 EKQERLAKLVGGVAVIHVGAATEEMKEKORVEDALNATRAAVEGIVPGGTA FVS 422  
DB 361 EKQERLAKLVGGVAVIKVGAATEEMKEKORVEDALNATRAAVEGIVAGGVALIRA 420  
QY 423 IKVLDDIKPADDDELAGLNIIRRSLEERPLQIAANAGYEGSVIVVEKREPKDGFNAAS 482  
DB 421 RTAIGLTGANADQAGIKVLRAMEERPLQIVTNGGEEASVVVAAGKNGVYNAAT 480  
QY 483 GEYEDLTKAGVIDPKVTRIALQNAASVASILLTTECAIAEKPKPKDMPMPGGMGMG 542  
DB 481 GEYVDMVEAGVDPKVTTRIALQNAASVAGLLTDDAAVAEL--PKEDA PMPGMPGGMG 538  
QY 543 GM 544

DB 539 GM 540

## RESULT 4

US-10-369-493-10014  
Sequence 10014, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 10014  
LENGTH: 547  
TYPE: PRT  
ORGANISM: magnetite-containing magnetic coccus  
US-10-369-493-10014

Query Match 68.6%; Score 1857; DB 15; Length 547;  
Best Local Similarity 69.5%; Pred. No. 3.2e-119;  
Matches 382; Conservative 72; Mismatches 90; Indels 6; Gaps 4;  
QY 1 MASKEILFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSFSGSPVITKDGVSVAKEI 60  
DB 1 MAAKEVKGFEAARAKMLNGVILANAVKVTLPKGRNVVLDKSGAPRMTKDGVSVAKEI 60  
QY 61 ELEDKFNENGAQWKEVAPKTSIADGDTTATVLAQAIYREGVLAAGRNPMATKRGI 120  
DB 61 ELEDKFNENGAQWKEVREVSSTADVAGDGTATVLAQAIYREGKMAVAAGRNPMATKRGI 120  
QY 121 DKAVAVATKELSDITKPTDQKEIAQVGTISANSDDTTIGNIIAEMAKVKGCVITVEEA 180  
DB 121 DLAVEAVVGLKIGISREVANSQEIQAQVGAISANSKDVGMIAEMDKVKEGVIIVVEEA 180  
QY 181 KGLETTLDVVEGKMFDRGYLSPYFTVNPENKVCLENDNPYILCNEKKITSMKMDLPILQEV 240  
DB 181 KGLETTLDVVEGQMFDRGYLSPYFTVNDKMLQVMPENLILLVEKKISNLQIILQILEGA 240

QY 241 AKVNRPLLIIEADVEGEALATLVNKLRGALQVAVKAPGFGERRKAMLEDAITLTGGEA 300  
DB 241 VQSRPLMIIEADVEGEALATLVNKLRGGLKVCVAVKAPGFGDRRKAMMEDIAITLTGGVL 300  
QY 301 IFEDRGKILNVSLSLGTAKRVVIDKENTTIVDAGKSEDIKARVKQIRAOIETSSDY 360  
DB 301 VSEDVGVKLENVNMDMLGMAKSIIVTKEDTTIDGAGDHEAIKARVQVQIRAOIETSSDY 360  
QY 361 DREKLOERLAKLVGGVAVIHVGAATEEMKEKORVEDALNATRAAAVEGIVPGGTA FV 420  
DB 361 DREKLOERLAKLVGGVAVIKVGGATEVEVKEKORVDDALHATRAAAVEGIVPGGVAL 420  
QY 421 RSLKV-LDDIKPADDDELAGLNIIRRSLEERPLQIAANAGYEGSVIVVEKREPKD-GFGF 478  
DB 421 RAEEASLTNIOGANDHQQOVGINIVTRALEEPLRIIASNAGAEGSVVVRVETKETNFGF 480  
QY 479 NAASGEYEDLTKAGVIDPKVTRIALQNAASVASILLTTECAIAEKPEPKDMP-MFGGG 537  
DB 481 NAATGYVEDLVASGVIDPAKVVRHALQNAASVAGLMTTTEAMVAELP--KDEPAMFGGD 537  
QY 538 MCGVGGMDGM 547  
DB 538 MCGVGGMGGM 547



```
Best Local Similarity 67.4%; Pred. No. 2.7e-118;
Matches 370; Conservative 74; Mismatches 99; Indels 7; Gaps 2;

Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSFSPVITKDGVSVAKEI 60
Db 1 MAARKVLPADAEARVIRVGVNVLANAVKVTLPKGRNVVILERSFGAPVITKDGVSVAKEI 60
Qy 61 ELEDPENMGAMQVKEVAPKTSIDTAGDGTATVLAQAIYREGVKLVAAAGRNPMIAIKRGI 120
Db 61 ELEDPENMGAMQVKEVAPKTSIDTAGDGTATVLAQAIYREGVKLVAAAGRNPMIAIKRGI 120
Qy 61 ELEDPENMGAMQVKEVAPKTSIDTAGDGTATVLAQAIYREGVKLVAAAGRNPMIAIKRGI 120
Db 61 ELEDPENMGAMQVKEVAPKTSIDTAGDGTATVLAQAIYREGVKLVAAAGRNPMIAIKRGI 120
Qy 121 DKAVAVTKELSDITKPTDQKEIAQVGTISANSDDTTIGNIABAMAKVGGVITVEEA 180
Db 121 DLAAEVVVKDLQAKAKKINTSEEVAVQVGTISANGERQIGLDIAEAMQVGVNEGVTVEEA 180
Qy 181 KGLETTLDVVEGKFPDRGYLSPYFVTNPKEVMVCELDNPNYILCNEKKITSMKMDLPILQOV 240
Db 181 KTAETELVEVGMQFDRGYLSPYFVTNPKEVMVCELDNPNYILCNEKKITSMKMDLPILQOV 240
Qy 241 AKVNRPLIIIAEDVEGELATLVNKLKRGALQVAVKAPGFGERRKAMLEIDIAITGTGEA 300
Db 241 VQTKPLVIAEDVEGELATLVNKLKRGALQVAVKAPGFGERRKAMLEIDIAITGTGTV 300
Qy 301 IFEDRGKLENVSLSSLTAKRVVVDKENTTIIVDGAGKSEDIKARVKQIRAIQIBETSSDY 360
Db 301 ISEDLGKLESVTLDMGLGSKKVSISKENTTIIVDGAGKQSDIEGRVAQIKAIQIBETSSDY 360
Qy 361 DREKQLERLAKLVGVAVIHVGAATETEMKEKKORVEDALNATRAAEEGIVPGGTAFAV 420
Db 361 DREKQLERLAKLVGVAVIRVGGSTEVEVKEKKRIDDALNATRAAEEGIVPGGTVALL 420
Qy 421 RSIKVLDDIKPADDELALNIIIRSLPEERQIAAANAGYEGSVIVKVRREP-KDGFQFN 479
Db 421 RSSFKI-TVKGVNDQDQAGINIVKALQSLVRQIAENAGDEASIVVGKILDKNEDNYGN 479
Qy 480 AASGEYEDLIKAGVIDPKVTRIALQNAASVASLLLTTECAIAIEKPEPKDMP-MPGGGM 538
Db 480 AQTGEYEDLIKAGVIDPKVTRIALQNAASVASLLLTTECAIAIEKPEPKDMP-MPGGGM 539
Qy 539 GGM 541
Db 540 GGM 542

RESULT 8
US-10-369-493-7917
; Sequence 7917, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7917
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Rhodobacter sphaeroides
US-10-369-493-7917

Query Match 67.9%; Score 1839.5; DB 15; Length 547;
Best Local Similarity 67.9%; Pred. No. 5.2e-118;
Matches 372; Conservative 70; Mismatches 103; Indels 3; Gaps 2;

Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSFSPVITKDGVSVAKEI 60
Db 1 MAARKVLPADAEARVIRVGVNVLANAVKVTLPKGRNVVILERSFGAPVITKDGVSVAKEI 60
Qy 61 ELEDPENMGAMQVKEVAPKTSIDTAGDGTATVLAQAIYREGVKLVAAAGRNPMIAIKRGI 120
Db 61 ELEDPENMGAMQVKEVAPKTSIDTAGDGTATVLAQAIYREGVKLVAAAGRNPMIAIKRGI 120
Qy 121 DKAVAVTKELSDITKPTDQKEIAQVGTISANSDDTTIGNIABAMAKVGGVITVEEA 180
Db 121 DLAAEVVVKDLQAKAKKINTSEEVAVQVGTISANGERQIGLDIAEAMQVGVNEGVTVEEA 180
Qy 181 KGLETTLDVVEGKFPDRGYLSPYFVTNPKEVMVCELDNPNYILCNEKKITSMKMDLPILQOV 240
Db 181 KTAETELVEVGMQFDRGYLSPYFVTNPKEVMVCELDNPNYILCNEKKITSMKMDLPILQOV 240
Qy 241 AKVNRPLIIIAEDVEGELATLVNKLKRGALQVAVKAPGFGERRKAMLEIDIAITGTGEA 300
Db 241 VQTKPLVIAEDVEGELATLVNKLKRGALQVAVKAPGFGERRKAMLEIDIAITGTGTV 300
Qy 301 IFEDRGKLENVSLSSLTAKRVVVDKENTTIIVDGAGKSEDIKARVKQIRAIQIBETSSDY 360
Db 301 ISEDLGKLESVTLDMGLGSKKVSISKENTTIIVDGAGKQSDIEGRVAQIKAIQIBETSSDY 360
Qy 361 DREKQLERLAKLVGVAVIHVGAATETEMKEKKORVEDALNATRAAEEGIVPGGTAFAV 420
Db 361 DREKQLERLAKLVGVAVIRVGGSTEVEVKEKKRIDDALNATRAAEEGIVPGGTVALL 420
Qy 421 RSIKVLDDIKPADDELALNIIIRSLPEERQIAAANAGYEGSVIVKVRREP-KDGFQFN 479
Db 421 RSSFKI-TVKGVNDQDQAGINIVKALQSLVRQIAENAGDEASIVVGKILDKNEDNYGN 479
Qy 480 AASGEYEDLIKAGVIDPKVTRIALQNAASVASLLLTTECAIAIEKPEPKDMP-MPGGGM 538
Db 480 AQTGEYEDLIKAGVIDPKVTRIALQNAASVASLLLTTECAIAIEKPEPKDMP-MPGGGM 539
Qy 539 GGM 541
Db 540 GGM 542

RESULT 7
US-10-369-493-11760
; Sequence 11760, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11760
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-11760

Query Match 68.0%; Score 1840.5; DB 15; Length 544;
Best Local Similarity 69.2%; Pred. No. 4.4e-118;
Matches 376; Conservative 69; Mismatches 95; Indels 3; Gaps 3;

Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSFSPVITKDGVSVAKEI 60
```





```
Qy 1 MASKEILFDKAREKLSRGVDKLANAVKTLGPKGRNVLEKSGSPVITKDGVSVAKEI 60
Db 1 MAADKIRFGEDARTMRVGRVGNVLANAVKATLGPGRNVLEKSGSPVITKDGVSVAKEI 60
Qy 61 ELEDKFENMGQAQVKEVAPKTSIDAGDGTATTATVLAQAIYREGVKLVAAAGNPNMAIKRGI 120
Db 61 ELADKFENMGQAQVKEVASRTNDNAGDGTATTATVLAQALIREGAKAVAAAGNPNMDLKRGI 120
Qy 121 DKAVAVTKELSDITKPTRDQKEIAQVGTISANSDDTTIGNIIAABAMAKVKGKGVITVEEA 180
Db 121 DQAVKAAVIELKNISKPTTDDKAIQVGTISANSDESIGNIIAABAMQKVGKGVITVEEG 180
Qy 181 KGLETTDVVVEGKMFDRGYLSPYFVTPNPEKMCVCELDNPNYILCNKKITSMKMDLPILQV 240
Db 181 SGLNELDVVEGQFDRGYLSPYFINNQSQSADLDDPFILLHDKKISNVRDLPLVLEGV 240
Qy 241 AKVNRPLLIITAEVDEGEALATLVNKLRGALQVAVKAPGFGERRKAMLEDIAITLGTGEA 300
Db 241 AKAGKPLLIIVAEVEGEALATLVNVTIRGIVKVVAVKAPGFGDRRKAMLEDMAVLGTGTV 300
Qy 301 IFEDRGIKLENVSLSSLTGAKRVVIDKENTTIIVDGAGKSEDIKARVKQIRAOIBETSSDY 360
Db 301 ISEEVGLALEKATIKDLGRAKKVQVSKENTTIIDGAGDSATIEARVQGIKTQIBEDTSSDY 360
Qy 361 DREKQLERLAKLVGGVAVIHVGAATETEMKEKKDRVEDALNATRAAAVEEGIVPGGGTAFV 420
Db 361 DREKQLERVAKLACGAVAVIKVGASTETEMKEKKARVEDALHATRAAAVEEGVPPGGVALV 420
Qy 421 RSIKVLDDIKPADDELALAGNIIRRSLEELPQIAALNAAGYEGSVIVBKEPKDGFNFNA 480
Db 421 RALVAVGNLTGANEDQTHGQIALRAMEAPLREIVANAGEEPSVILNKVKEGTGNYGNA 480
Qy 481 ASGEYEDLIKAGVIDPKKVTTRIALQNAASVASLLLTTECAIAEKPEPKDMP-MP-CGGM 538
Db 481 ANGEFGDMVBEFGILDPTKVTRSALQNAASITAGLMITTEAMVAD--APKDEPAMPAGGGM 538
Qy 539 GGMGGM 544
Db 539 GGMGGM 544
RESULT 13
US-10-369-493-16217
; Sequence 16217, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 16217
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-16217
Query Match 67.7%; Score 1833; DB 15; Length 544;
Best Local Similarity 66.8%; Pred. No. 1.4e-117;
Matches 365; Conservative 82; Mismatches 95; Indels 4; Gaps 3;
```

```
Qy 61 ELEDKFENMGQAQVKEVAPKTSIDAGDGTATTATVLAQAIYREGVKLVAAAGNPNMAIKRGI 120
Db 61 ELADKFENMGQAQVKEVASRTNDNAGDGTATTATVLAQALIREGAKAVAAAGNPNMDLKRGI 120
Qy 121 DKAVAVTKELSDITKPTRDQKEIAQVGTISANSDDTTIGNIIAABAMAKVKGKGVITVEEA 180
Db 121 DQAVKAAVIELKNISKPTTDDKAIQVGTISANSDESIGNIIAABAMQKVGKGVITVEEG 180
Qy 181 KGLETTDVVVEGKMFDRGYLSPYFVTPNPEKMCVCELDNPNYILCNKKITSMKMDLPILQV 240
Db 181 SGLNELDVVEGQFDRGYLSPYFINNQSQSADLDDPFILLHDKKISNVRDLPLVLEGV 240
Qy 241 AKVNRPLLIITAEVDEGEALATLVNKLRGALQVAVKAPGFGERRKAMLEDIAITLGTGEA 300
Db 241 AKAGKPLLIIVAEVEGEALATLVNVTIRGIVKVVAVKAPGFGDRRKAMLEDMAVLGTGTV 300
Qy 301 IFEDRGIKLENVSLSSLTGAKRVVIDKENTTIIVDGAGKSEDIKARVKQIRAOIBETSSDY 360
Db 301 ISEEVGLALEKATIKDLGRAKKVQVSKENTTIIDGAGDSATIEARVQGIKTQIBEDTSSDY 360
Qy 361 DREKQLERLAKLVGGVAVIHVGAATETEMKEKKDRVEDALNATRAAAVEEGIVPGGGTAFV 420
Db 361 DREKQLERVAKLACGAVAVIKVGASTETEMKEKKARVEDALHATRAAAVEEGVPPGGVALV 420
Qy 421 RSIKVLDDIKPADDELALAGNIIRRSLEELPQIAALNAAGYEGSVIVBKEPKDGFNFNA 480
Db 421 RALVAVGNLTGANEDQTHGQIALRAMEAPLREIVANAGEEPSVILNKVKEGTGNYGNA 480
Qy 481 ASGEYEDLIKAGVIDPKKVTTRIALQNAASVASLLLTTECAIAEKPEPKDMP-MP-CGGM 538
Db 481 ANGEFGDMVBEFGILDPTKVTRSALQNAASITAGLMITTEAMVAD--APKDEPAMPAGGGM 538
Qy 539 GGMGGM 544
Db 539 GGMGGM 544
RESULT 14
US-10-282-122A-63193
; Sequence 63193, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-09
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
```

```
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 63193
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-10-369-493-15265

Query Match      67.7%; Score 1832; DB 15; Length 545;
Best Local Similarity 65.9%; Pred. No. 1.7e-117;
Matches 359; Conservative 84; Mismatches 100; Indels 2; Gaps 1;

Qy 3 SKETLFDKAREKLSRGVDKLANAVKTLGPKGRNVVIEKSGFSPVITKDGVSVAKEI 62
Db 2 AKDVSFGSNAREKIMDGVNILLADAVKVTGPKGRNVVIDKSGFAPITTKDGVSVAKEI 61

Qy 63 EDKFNMGAAQVKEVAPKTSIDAGDGTATTATVLAQAIYREGVKLVAAAGNPMATKRGIDK 122
Db 62 EDKFNMGAAQVREKASNTDNDVAGDGTATTATVLAQAIYREGVKLVAAAGNPMATKRGIDK 121

Qy 123 AVAVATKELSDITKPTDQKEIAQVGTISANSDDTTIGNIIAEMAKVKGKGVITVEBAK 182
Db 122 AVRAVEEIRAISTPANDHKAIAQVGSISANSDDTTIGELISKANETVGKQGVITVEBSG 181

Qy 183 LETTLQVVEGKMFDRGYLSPYFVTNPERKMWCELDNPIYLCNEKKITSMKMDLPILQV 242
Db 182 PEDALEVVEGQMFDRGYLSPYFANKQDLSLTCFDFNPFIILLVDDKISNIREIVPLEKVMQ 241

Qy 243 VNRPLIIIAEDVEGEALATLVNKLQVAVKAPGFGERRKAMLEDIAITLTGGEAIF 302
Db 242 TSREPLIIIAEDVEGEALATLVNKLQVAVKAPGFGERRKAMLEDIAITLTGGEVVIS 301

Qy 303 EDKGIKLENVSLSLGTAKRVVIDKENTTIVDGGKSEDIKARVKQIRAQIETSSDYDR 362
Db 302 EEVGLSETAIEHLGTAKVTIGKENTVIVDGGKSEDIKARVESIRROVEESTSDYDK 361

Qy 363 EKLQERLAKVGGVAVIHVGAATETEMKEKDRVEDALNATRAAEEGIVPGGTAFVRS 422
Db 362 EKLQERLAKVGGVAVIKVGAATETEMKEKDRVEDALNATRAAEEGIVPGGTAFVRS 421

Qy 423 IKVLDDIKPADDDDELAGNIIRRSLEELPLQIAANAGYEGSIVVEKREPDKDGFENAS 482
Db 422 LSALSDUKDNDQAGNIIRRAWEAPLQIVSNAGDEASVIVNEKNGSGNYGNAAS 481

Qy 483 GEYEDLIKAGVIDPKVTRIALQNAASVASILLTTECAIAEKPPKMDMPMPGGMGMG 542
Db 482 GEYDMLGEMILDPAKVTRSALEHAASVAGMLTTEVMTDKPAP--EAPMPAGGMGMG 539

Qy 543 GMDGM 547
Db 540 GMDGM 544
```

## RESULT 15

```
US-10-369-493-15265
; Sequence 15265, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
```

```
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15265
; LENGTH: 543
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-15265
```

```
Query Match      67.5%; Score 1828.5; DB 15; Length 543;
Best Local Similarity 68.9%; Pred. No. 2.9e-117;
Matches 374; Conservative 70; Mismatches 96; Indels 3; Gaps 3;

Qy 1 MASKEILFDKAREKLSRGVDKLANAVKTLGPKGRNVVIEKSGFSPVITKDGVSVAKEI 60
Db 1 MAAKEVPGASAREKMLKGVDDILLADAVKVTGPKGRNVVIDKSGFAPITTKDGVSVAKEI 60

Qy 61 ELEDKFNMGAAQVKEVAPKTSIDAGDGTATTATVLAQAIYREGVKLVAAAGNPMATKRG 120
Db 61 ELEDKFNMGAAQVREKASNTDNDVAGDGTATTATVLAQAIYREGKAVAAAGNPMATKRG 120

Qy 121 DKAVVATKELSDITKPTDQKEIAQVGTISANSDDTTIGNIIAEMAKVKGKGVITVEEA 180
Db 121 DLAAEVVVKDLQAKAKINTSEEAQVGTISANGEROIGLDIAEMQVNGEGVITVEEA 180

Qy 181 KGLETTLDVVEGKMFDRGYLSPYFVTNPERKMWCELDNPIYLCNEKKITSMKMDLPILQV 240
Db 181 KTAETELVVEGQMFDRGYLSPYFVTNPERKMWADLEADAYILLHEKLSNLAQMLPVLEAV 240

Qy 241 AKVNRPLIIIAEDVEGEALATLVNKLQVAVKAPGFGERRKAMLEDIAITLTGGEA 300
Db 241 VQTGKPLVIAEDVEGEALATLVNKLQVAVKAPGFGERRKAMLEDIAITLTGTV 300

Qy 301 IFEDRGIKLENVSLSLGTAKRVVIDKENTTIVDGGKSEDIKARVKQIRAQIETSSDY 360
Db 301 ISEDLGKLSVSLDMLGKSKVSIKENTTIVDGGKSEDIKARVQAIKQIETSSDY 360

Qy 361 DREKLQERLAKVGGVAVIHVGAATETEMKEKDRVEDALNATRAAEEGIVPGGTAFV 420
Db 361 DREKLQERLAKVGGVAVIRVGGSTEVEVEKEDRIDALNATRAAEEGIVPGGVALL 420

Qy 421 RSKVLDDIKPADDDDELAGNIIRRSLEELPLQIAANAGYEGSIVVEKREP-KDGFEN 479
Db 421 RSSTKI-TVKGVNDQDQAGINIVRKALQSLVRQIAENAGDEASIVWGKILDKNEDNYN 479

Qy 480 AASGEYEDLIKAGVIDPKVTRIALQNAASVASILLTTECAIAEKPPKMDMP-MPGGM 538
Db 480 AQTGEYDGLIALGIVDPKVVTRIALQNAASVASVLIITEAMIAELPKKESAMPQPGGM 539

Qy 539 GGM 541
Db 540 GGM 542
```

Search completed: April 12, 2005, 09:51:10  
Job time : 138 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 09:35:27 ; Search time 43 seconds  
(without alignments)  
951.342 Million cell updates/sec

Title: US-09-077-574A-2

Perfect score: 2708

Sequence: 1 MASKEILFDKAREKLSRGV.....KDMPPGGMGGMGGMDGMY 548

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1874.5	69.2	547	4	US-09-902-540-13219
2	1844	68.1	548	2	US-08-467-822-31
3	1844	68.1	548	3	US-08-432-697-31
4	1844	68.1	548	3	US-08-466-248-31
5	1832	67.7	546	4	US-09-540-236-2574
6	1820.5	67.2	551	4	US-09-252-991A-24209
7	1768.5	65.3	550	4	US-09-328-352-5508
8	1765	65.2	557	4	US-09-543-681A-6606
9	1746	64.5	545	4	US-09-544-683-2
10	1746	64.5	545	4	US-10-192-419-2
11	1741	64.3	548	2	US-08-467-822-32
12	1741	64.3	548	3	US-09-472-971-3
13	1741	64.3	548	3	US-08-432-697-32
14	1741	64.3	548	3	US-08-466-248-32
15	1723.5	63.6	546	3	US-08-470-260-6
16	1723.5	63.6	546	3	US-08-471-491-6
17	1723.5	63.6	546	3	US-08-466-662-6
18	1723.5	63.6	546	4	US-08-256-847C-1
19	1723.5	63.6	546	4	US-08-256-847C-7
20	1721	63.6	544	2	US-08-467-822-33
21	1721	63.6	544	3	US-08-432-697-33
22	1721	63.6	544	3	US-08-466-248-33
23	1716.5	63.4	547	3	US-08-461-722-2
24	1716.5	63.4	547	3	US-08-336-251-2
25	1716.5	63.4	547	5	PCT-US94-06362-2
26	1716.5	63.4	547	5	PCT-US94-06362-2
27	1710	63.1	544	4	US-09-809-745-3

Sequence 136, App  
Sequence 153, App  
Sequence 30, Appl  
Sequence 30, Appl  
Sequence 30, Appl  
Sequence 30, Appl  
Sequence 51, Appl  
Sequence 51, Appl  
Sequence 3685, Ap  
Sequence 4559, Ap  
Sequence 790, App  
Sequence 160, App  
Sequence 160, App  
Sequence 160, App  
Sequence 160, App  
Sequence 160, App  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 6, Appli

28 1710 63.1 555 4 US-09-438-185A-136  
29 1708 63.1 544 4 US-09-198-452A-153  
30 1702 62.9 545 2 US-08-467-822-30  
31 1702 62.9 545 3 US-08-432-697-30  
32 1702 62.9 545 3 US-08-466-248-30  
33 1683.5 62.2 641 4 US-09-613-303-51  
34 1683.5 62.2 641 4 US-10-267-311-51  
35 1674.5 61.8 540 4 US-09-583-110-3685  
36 1671 61.7 545 4 US-09-134-000C-4559  
37 1624.5 60.0 539 4 US-09-710-279-790  
38 1597 59.0 541 2 US-08-997-080-160  
39 1597 59.0 541 2 US-08-997-362-160  
40 1597 59.0 541 3 US-09-095-855-160  
41 1597 59.0 541 3 US-09-324-542-160  
42 1597 59.0 541 4 US-09-205-426-160  
43 1595.5 58.9 540 4 US-09-613-303-4  
44 1595.5 58.9 540 4 US-10-267-311-4  
45 1595.5 58.9 540 4 US-09-847-637B-6

#### ALIGNMENTS

#### RESULT 1

US-09-902-540-13219  
; Sequence 13219, Application US/09902540  
; Patent No. 6833447

GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Wiegand, Roger C.

TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

FILE REFERENCES: 38-10(15849)B

CURRENT APPLICATION NUMBER: US/09/902,540

CURRENT FILING DATE: 2001-07-10

PRIOR APPLICATION NUMBER: 60/217,883

PRIOR FILING DATE: 2000-07-10

NUMBER OF SEQ ID NOS: 16825

SEQ ID NO 13219

LENGTH: 547

TYPE: PRT

ORGANISM: Myxococcus xanthus

US-09-902-540-13219

Query Match 69.2% Score 1874.5; DB 4; Length 547;

Best Local Similarity 69.2% Pred. No. 2.7e-155;

Matches 379; Conservative 69; Mismatches 99; Indels 1; Gaps 1;

Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVTLPKGRNVVIEKSPGSPVITKDGVSVAKEI 60  
Db 1 MAAKEIFPHQSAREAILRGVRTLSDAVAVTLGPGRNVVIEKSPGSPVITKDGVTVAKEI 60  
Qy 61 ELEDKFNMAQNVKEVAPKTSIAGDGTATTATVLAQAIYREGVKVLAAGNPMAIKRGI 120  
Db 61 DLENKFNMAQNVKEVAPKTSIAGDGTATTATVLAQAIYREGVKVLAAGNPMAIKRGI 120  
Qy 121 DKAVAVVTKELSDITKPTQKEIAQVGTISANSDDTTIGNIIAEMAKVKGKGVITVEEA 180  
Db 121 DKAVEVVVGLKSLSKPTADKKAITQVGTISANGDEITGAIADAMEKVGKGVITVEEA 180  
Qy 181 KGLETTLDVVEGMMKFRGYLSPYFVTNPERKVCLENDPNYILCNKKITSMKMDLPILQV 240  
Db 181 KGLETTLDVVEGMMKFRGYLSPYFVTNPERKVCLENDPNYILCNKKITSMKMDLPILQV 240  
Qy 241 AKVNRPLIIAEDVVEGALATLVNKLGLAQVAVKAPGFGERRKAMLEDAIILTGEA 300  
Db 241 ARSKPLIIAEDVVEGALATLVNKLGLAQVAVKAPGFGERRKAMLEDAIILTGEA 300  
Qy 301 IFEDRGIKLENVSLSSIGTAKRVVIDKENTIVDAGKSEDIKARVKQIRAIQIETSSDY 360  
Db 301 VSEDLGHKFTLTITDLGRKRVTVDKNTIVDVGVTGKAAIEGRIKLIRTIQIDSVTSDY 360



```
;
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-432-697-31

Query Match 68.1%; Score 1844; DB 3; Length 548;
Best Local Similarity 67.2%; Pred. No. 1.2e-152;
Matches 366; Conservative 81; Mismatches 96; Indels 2; Gaps 1;

QY 3 SKEILFADAKAREKLSRGVDKLANAVKVTGLPGKGRNVVIEKSFSGSPVITKDGVSVAKEIEL 62
DB 2 AKELRFDDARLQMLAGVNALADAVQVTMGPRGRNVVLEKSYGAPTPTKDGVSVAKEIEF 61

QY 63 EDFENMGQAVKVEAPKTSDIAGDGTATTATVLAQAIYREGVKLVAAAGRNPMATKRGIDK 122
DB 62 EHRFPMNGQAVKVEASKTSDTAGDGTATTATVLAQAIYREGVKLVAAAGRNPMATKRGIDK 121

QY 123 AVAVATKELSDITKPTDRDQKEIAQVGTISANSDDTTIGNIIAEMAKYKGGVITVEBAKG 182
DB 122 AVLATVTKLQAMSKPKCKSKAIAQVGTISANSDEAIGAIIEAMEKVGKGVITVEDGNG 181

QY 183 LETTLDVVEGMKFDGRGYLSFYFTNPEKVMCELDNPNVILCNEKKTSMKMDLPILEOVAK 242
DB 182 LENELYVVEGMQFDRGYISFYFTNPNQNMSCLEHPFILLVDDKKVSSIREMLSVLEGVAK 241

QY 243 VNRPLLIADVEGEALATLVNKLKALGVAVKAPFGGERRKAMLEDAITLTGGGAIF 302
DB 242 SGRPLLIADIEGEALATLVNMMRGIVKVCAPFGFGDRRKAMLIQDAILTKGQVIS 301

QY 303 EDGKIKLENVSLSGTAKRWIDKENTTIVDGAAGKSEDIKARVKQIRAOIETSSDYDR 362
DB 302 EBIKSLLEGATLEDIGSAKRIVVTENTTIIDGEGKATEINARIAQIRAOIETSSDYDR 361

QY 363 EKQLERLAKLVGGVAVTHVGAATETEMKEDRVEDALNATRAAVEGIVPGGTAFAVRS 422
DB 362 EKQLERLAKLVGGVAVTHVGAATEVENKEKARVEDALNATRAAVEGIVAGGVALIRA 421

QY 423 IKVLDDIKPADDDELAGLNIIRSLPELRQIAANAGYEGSIVVEKVRPEKDGFGFNAAS 482
DB 422 QKALDSLKGNDQNMGINILRRATIESPMRQIVTNAGYEASVVVNVKVAEKKDNYGFNAAT 481

QY 483 GEYEDLLKAGVIDPKKVTALQNAASVASLILTECAIAEKPEPKDMPMPGCGMG 542
DB 482 GEYEDLVEMGILDPKVTALQNAASVASLILTECAIAEKPEPKDMPMPGCGMG 542

QY 543 GMDGM 547
DB 540 GMDGM 544

RESULT 4
US-08-466-248-31
; Sequence 31, Application US/08466248
; Patent No. 6258359
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
```

```
;
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,248
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-466-248-31

Query Match 68.1%; Score 1844; DB 3; Length 548;
Best Local Similarity 67.2%; Pred. No. 1.2e-152;
Matches 366; Conservative 81; Mismatches 96; Indels 2; Gaps 1;

QY 3 SKEILFADAKAREKLSRGVDKLANAVKVTGLPGKGRNVVIEKSFSGSPVITKDGVSVAKEIEL 62
DB 2 AKELRFDDARLQMLAGVNALADAVQVTMGPRGRNVVLEKSYGAPTPTKDGVSVAKEIEF 61

QY 63 EDFENMGQAVKVEAPKTSDIAGDGTATTATVLAQAIYREGVKLVAAAGRNPMATKRGIDK 122
DB 62 EHRFPMNGQAVKVEASKTSDTAGDGTATTATVLAQAIYREGVKLVAAAGRNPMATKRGIDK 121

QY 123 AVAVATKELSDITKPTDRDQKEIAQVGTISANSDDTTIGNIIAEMAKYKGGVITVEBAKG 182
DB 122 AVLATVTKLQAMSKPKCKSKAIAQVGTISANSDEAIGAIIEAMEKVGKGVITVEDGNG 181

QY 183 LETTLDVVEGMKFDGRGYLSFYFTNPEKVMCELDNPNVILCNEKKTSMKMDLPILEOVAK 242
DB 182 LENELYVVEGMQFDRGYISFYFTNPNQNMSCLEHPFILLVDDKKVSSIREMLSVLEGVAK 241

QY 243 VNRPLLIADVEGEALATLVNKLKALGVAVKAPFGGERRKAMLEDAITLTGGGAIF 302
DB 242 SGRPLLIADIEGEALATLVNMMRGIVKVCAPFGFGDRRKAMLIQDAILTKGQVIS 301

QY 303 EDGKIKLENVSLSGTAKRWIDKENTTIVDGAAGKSEDIKARVKQIRAOIETSSDYDR 362
DB 302 EBIKSLLEGATLEDIGSAKRIVVTENTTIIDGEGKATEINARIAQIRAOIETSSDYDR 361
```





```
; LENGTH: 545
; TYPE: PR1
; ORGANISM: Campylobacter jejuni
US-09-544-683-2

Query Match      64.5%; Score 1746; DB 4; Length 545;
Best Local Similarity 64.5%; Pred. No. 4.6e-144;
Matches 353; Conservative 80; Mismatches 108; Indels 6; Gaps 4;

Qy  3 SKEILFADAKAREKLSRGVDKLANAVKVTGLPGKRNVIIEKSFSGPVITKDGVSVAKEIEL 62
Db  2 AKEIIFDEARNKLYEGVKLNDVAVKVTGMPGRNVLIQKSFGPSITKDGVSVAKEVEL 61

Qy  63 EDKPENNGAOMVEKVA PKTSDIAGDGTATTATVLAQAIIYREGVKLVAAAGRNPMATIKRGIDK 122
Db  62 KDSLENNGASLVREVASKTADQAGDGTATTATVLAHAIFKEGLRNITAGANPIEVKRGMDK 121

Qy  123 AVAVTKELSDITKPTRDQKEIAQVGTISANSDDTTIGNIIAEMAKVKGKGVITVEBAKG 182
Db  122 ACEAIVAEKLSREVKDKKEIAQVATISANSDEKIGNLIADAMEKVGKDGVIIVEBPKS 181

Qy  183 LETTLDVVEGKDFRGYLSPYFVTNPEKMYCELDNPNVILCNEKITSKMDMLPILEQVAK 242
Db  182 INDELNVVEGQDFRGYLSPYFITNAEKMTVELSSPYILLFDKKITNLKOLLPVLEQIQK 241

Qy  243 VNRPLIIAEDVEGEALATLVNKLRGALOVVAVKAPGFGERRKAMLEDAITLTGGEAIF 302
Db  242 TGRPLIIAEDIEGEALATLVNKLRGVLNLSAVKAPFGDGRKAMLEDAITLTGGEVIS 301

Qy  303 EDGKILENVSLSSLGTAKEVVIDKENTTIVDGAGKSEDIKARVKQIRAQIEETSSDYDR 362
Db  302 EELGRTLESATIQDLQOASSVIDKNTTIVNGAGEKANIDARVNQIKAQIAETTSYDR 361

Qy  363 EKLOERLAKLVGGVAVIHVGAATEMKEKKORVEDALNATRAAVERGIVPGGTAFAVRS 422
Db  362 EKLOERLAKLVGGVAVIKVGATTETEMKEKKDRVDDALSATKAVERGIVIGGGAALIK 421

Qy  423 -IKVLDIKPADDELALAGLNIIRSLPELRQIAANAGYEGSIIVVEKVRPKD-GFGFNA 480
Db  422 KAKIKLDLQ--GDEAIGAAIVERALRAPLQIAENAGFDAGVVVNSVENAKDENTGFDA 478

Qy  481 ASGYEDLIKAGVIDPKKVTRIALQNAASVASLLLTTECAIAEKPEPKDMPMPGGMG 540
Db  479 AKGEYVNMLESIGIIDPVKVERVALLNNAVSVASMLLTTEATISEIKEDKPTMP-DMSGMG 537

Qy  541 MGNMGM 547
Db  538 MGNMGM 544

RESULT 10
US-10-192-419-2
; Sequence 2, Application US/10192419
; Patent No. 6790446
; GENERAL INFORMATION:
; APPLICANT: JACOBS, ANTONIUS A.C.
; APPLICANT: VAN DEN BOSCH, JOHANNES F.
; APPLICANT: NUIJTEN, PETRUS J.M.
; TITLE OF INVENTION: CAMPYLOBACTER VACCINE
; FILE REFERENCE: JACOBS
; CURRENT APPLICATION NUMBER: US/10/192,419
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US/09/544,683
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: EP99201086.8
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 545
; TYPE: PR1
; ORGANISM: Campylobacter jejuni
US-10-192-419-2

Query Match      64.5%; Score 1746; DB 4; Length 545;
Best Local Similarity 64.5%; Pred. No. 4.6e-144;
Matches 353; Conservative 80; Mismatches 108; Indels 6; Gaps 4;

Qy  3 SKEILFADAKAREKLSRGVDKLANAVKVTGLPGKRNVIIEKSFSGPVITKDGVSVAKEIEL 62
Db  2 AKEIIFDEARNKLYEGVKLNDVAVKVTGMPGRNVLIQKSFGPSITKDGVSVAKEVEL 61

Qy  63 EDKPENNGAOMVEKVA PKTSDIAGDGTATTATVLAQAIIYREGVKLVAAAGRNPMATIKRGIDK 122
Db  62 KDSLENNGASLVREVASKTADQAGDGTATTATVLAHAIFKEGLRNITAGANPIEVKRGMDK 121

Qy  123 AVAVTKELSDITKPTRDQKEIAQVGTISANSDDTTIGNIIAEMAKVKGKGVITVEBAKG 182
Db  122 ACEAIVAEKLSREVKDKKEIAQVATISANSDEKIGNLIADAMEKVGKDGVIIVEBPKS 181

Qy  183 LETTLDVVEGKDFRGYLSPYFVTNPEKMYCELDNPNVILCNEKITSKMDMLPILEQVAK 242
Db  182 INDELNVVEGQDFRGYLSPYFITNAEKMTVELSSPYILLFDKKITNLKOLLPVLEQIQK 241

Qy  243 VNRPLIIAEDVEGEALATLVNKLRGALOVVAVKAPGFGERRKAMLEDAITLTGGEAIF 302
Db  242 TGRPLIIAEDIEGEALATLVNKLRGVLNLSAVKAPFGDGRKAMLEDAITLTGGEVIS 301

Qy  303 EDGKILENVSLSSLGTAKEVVIDKENTTIVDGAGKSEDIKARVKQIRAQIEETSSDYDR 362
Db  302 EELGRTLESATIQDLQOASSVIDKNTTIVNGAGEKANIDARVNQIKAQIAETTSYDR 361

Qy  363 EKLOERLAKLVGGVAVIHVGAATEMKEKKORVEDALNATRAAVERGIVPGGTAFAVRS 422
Db  362 EKLOERLAKLVGGVAVIKVGATTETEMKEKKDRVDDALSATKAVERGIVIGGGAALIK 421

Qy  423 -IKVLDIKPADDELALAGLNIIRSLPELRQIAANAGYEGSIIVVEKVRPKD-GFGFNA 480
Db  422 KAKIKLDLQ--GDEAIGAAIVERALRAPLQIAENAGFDAGVVVNSVENAKDENTGFDA 478

Qy  481 ASGYEDLIKAGVIDPKKVTRIALQNAASVASLLLTTECAIAEKPEPKDMPMPGGMG 540
Db  479 AKGEYVNMLESIGIIDPVKVERVALLNNAVSVASMLLTTEATISEIKEDKPTMP-DMSGMG 537

Qy  541 MGNMGM 547
Db  538 MGNMGM 544

RESULT 11
US-08-467-822-32
; Sequence 32, Application US/08467822
; Patent No. 5843460
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,822  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/447,177  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/432,697  
; FILING DATE: 02-MAY-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 03495.0137-02000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 408-4000  
; TELEFAX: (202) 408-4400  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 548 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-467-822-32

Query Match 64.3%; Score 1741; DB 2; Length 548;  
Best Local Similarity 64.3%; Pred. No. 1.3e-143; Indels 4; Gaps 2;  
Matches 353; Conservative 74; Mismatches 118;  
  
Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVTGLPKGRNVVIEKSPGSPVITKDGVSVAKEI 60  
Db 1 MAADKVFNGDARVKMLRGVNVLDADVKTGLPKGRNVVLDKSPGAPITTKDGVSVAKEI 60  
  
Qy 61 ELEDKFNMGAMQVKEVAPKTSIDAGDGTATTATVLAQAIYREGVKLVAAGRNPMAIKRGI 120  
Db 61 ELEDKFNMGAMQVKEVASKANDAGDGTATTATVLAQAIITEGLKVAAGNPNMDLKRGI 60  
  
Qy 121 DKAVVAVTKELSDITTKPTDQKEIAQVGTISANSDDTTIGNIIAEAMAKVKGKGGVITVVEEA 180  
Db 121 DKAVTAAVEELKALSVPCSDSKALAQVGTISANSDETGVKLIASAMDKVKGKGGVITVEDG 180  
  
Qy 181 KGLETTLDVVEGKMFDRGYLSPYFVTNPKNMVCBLDNPYILCNCKITSMKMDLPVLEAV 240  
Db 181 TGLQDELVDVVEGMQFDRGYLSPYFINKPETGAVELESFPILLADKKISNIREMLPVLEAV 240  
  
Qy 241 AKVNRPLIIIAEDVEGEALATLVNKLRGALQVAVKAPGFGERRKAMLEDIALITGGEA 300  
Db 241 AKAGKPLIIIAEDVEGEALATAVNTIRGIVKVAAVKAPGFGDRRKAMLQDIATLTGTV 300  
  
Qy 301 IFEDRGIKLENVSLSSLTAKRVVIDKENTTIVDGAGKSEDIKARVKQIRAOIEETSSDY 360  
Db 301 ISEIGHLEKATLEDLGQAKRVVINKDTHIIIDVGVEEAAIQGRVAQIROQIEATSDY 360  
  
Qy 361 DREKLQERLAKLVGGVAVIHVGAATETEMKEKKDRVEDALNATRAAAVEEGIVPGGTAFAV 420  
Db 361 DREKLQERVAKLVGGVAVIKVGAATEVEMKEKKARVEDALHATRAAAVEEGVAVGGVALLI 420  
  
Qy 421 RSIKVLDDIKPADDELALNIIIRSEEPRLQIAANAGYEGSVVEKVPKDGFGFENA 480  
Db 421 RVASKLADLRQNEQDQNVGKVALRAMEAPLRQIVLNCGEEPSVVANTVKGDDGNGYNA 480  
  
Qy 481 ASGEYEDLIKAGVIDPKVTTRIALQNAASVASLILLTTTECAIAEKPEPKDMPMPG--GGM 538  
Db 481 ATEYGNWIDMGIIDPTKVTRSALQYAAASVAGLITTECMWTDL--PKNDADLGAAGGM 538  
  
Qy 539 GGMGMDGM 547  
Db 539 GGMGMDGM 547

RESULT 13  
US-08-432-697-32  
; Sequence 32, Application US/08432697  
; Patent No. 6248330  
; GENERAL INFORMATION:  
; APPLICANT: Labigne, Agnes

RESULT 12  
US-09-472-971-3  
; Sequence 3, Application US/09472971  
; Patent No. 6197547  
; GENERAL INFORMATION:  
; APPLICANT: SOGO, Kazuyo  
; APPLICANT: YANAGI, Hideki  
; APPLICANT: YURA, Takashi  
; TITLE OF INVENTION: TRIGGER FACTOR EXPRESSION PLASMIDS  
; FILE REFERENCE: 1422-409P  
; CURRENT APPLICATION NUMBER: US/09/472,971  
; CURRENT FILING DATE: 1999-12-28  
; EARLIER APPLICATION NUMBER: JP10-372965  
; EARLIER FILING DATE: 1998-12-28  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 548  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
; US-09-472-971-3

Query Match 64.3%; Score 1741; DB 3; Length 548;  
Best Local Similarity 64.3%; Pred. No. 1.3e-143; Indels 4; Gaps 2;  
Matches 353; Conservative 74; Mismatches 118;  
  
Qy 1 MASKEILFDKAREKLSRGVDKLANAVKVTGLPKGRNVVIEKSPGSPVITKDGVSVAKEI 60  
Db 1 MAADKVFNGDARVKMLRGVNVLDADVKTGLPKGRNVVLDKSPGAPITTKDGVSVAKEI 60  
  
Qy 61 ELEDKFNMGAMQVKEVAPKTSIDAGDGTATTATVLAQAIYREGVKLVAAGRNPMAIKRGI 120  
Db 61 ELEDKFNMGAMQVKEVASKANDAGDGTATTATVLAQAIITEGLKVAAGNPNMDLKRGI 120  
  
Qy 121 DKAVVAVTKELSDITTKPTDQKEIAQVGTISANSDDTTIGNIIAEAMAKVKGKGGVITVVEEA 180  
Db 121 DKAVTAAVEELKALSVPCSDSKALAQVGTISANSDETGVKLIASAMDKVKGKGGVITVEDG 180  
  
Qy 181 KGLETTLDVVEGKMFDRGYLSPYFVTNPKNMVCBLDNPYILCNCKITSMKMDLPVLEAV 240  
Db 181 TGLQDELVDVVEGMQFDRGYLSPYFINKPETGAVELESFPILLADKKISNIREMLPVLEAV 240  
  
Qy 241 AKVNRPLIIIAEDVEGEALATLVNKLRGALQVAVKAPGFGERRKAMLEDIALITGGEA 300  
Db 241 AKAGKPLIIIAEDVEGEALATAVNTIRGIVKVAAVKAPGFGDRRKAMLQDIATLTGTV 300  
  
Qy 301 IFEDRGIKLENVSLSSLTAKRVVIDKENTTIVDGAGKSEDIKARVKQIRAOIEETSSDY 360  
Db 301 ISEIGHLEKATLEDLGQAKRVVINKDTHIIIDVGVEEAAIQGRVAQIROQIEATSDY 360  
  
Qy 361 DREKLQERLAKLVGGVAVIHVGAATETEMKEKKDRVEDALNATRAAAVEEGIVPGGTAFAV 420  
Db 361 DREKLQERVAKLVGGVAVIKVGAATEVEMKEKKARVEDALHATRAAAVEEGVAVGGVALLI 420  
  
Qy 421 RSIKVLDDIKPADDELALNIIIRSEEPRLQIAANAGYEGSVVEKVPKDGFGFENA 480  
Db 421 RVASKLADLRQNEQDQNVGKVALRAMEAPLRQIVLNCGEEPSVVANTVKGDDGNGYNA 480  
  
Qy 481 ASGEYEDLIKAGVIDPKVTTRIALQNAASVASLILLTTTECAIAEKPEPKDMPMPG--GGM 538  
Db 481 ATEYGNWIDMGIIDPTKVTRSALQYAAASVAGLITTECMWTDL--PKNDADLGAAGGM 538  
  
Qy 539 GGMGMDGM 547  
Db 539 GGMGMDGM 547

RESULT 13  
US-08-432-697-32  
; Sequence 32, Application US/08432697  
; Patent No. 6248330  
; GENERAL INFORMATION:  
; APPLICANT: Labigne, Agnes





Db 1 MAADKVPNGDARVKMLRGVNLDAVKVTIGPGRNVVLDKSGAPITTKDGVSVAREI 60  
Qy 61 ELEDKFNMGQMKVEKAPKTSIAGDGTATVLAQAIYREGVKLVAAAGRNPMIAIKRGI 120  
Db 61 ELEDKFNMGQMKVEKASKANDAGDGTATVLAQAIITEGLKAVAGNPNPMDLKRGI 120  
Qy 121 DKAVAVTKELSDITKPRDQKEIAQVGTISANSDDTTGNIIAAMAKVKGKGVITVEEA 180  
Db 121 DKAVTAAVEELKALSVPCSDSKAIAQVGTISANSDETGVKLIABAMDKVKGKGVITVEDG 180  
Qy 181 KGLTTLDVVGKMFDRGYLSPYFVTNPEKVMVCELDNPEYILCNEKKITSMKMDLPILBOV 240  
Db 181 TGLQELDVVGKMFDRGYLSPYFINKPETGAVELESFPIILLADKKISNIREMLPVLAV 240  
Qy 241 AKVNRPLLIIEADVEGEALATLVNKLKRGALQVAVKAPGFGERRKAMLEDAIILTGEA 300  
Db 241 AKAGKPLLIIEADVEGEALATAVNTIRGIVKVAAVKAPGFGDRRKAMLDIATLTGGTV 300  
Qy 301 IFBDRGKLENVSLSSIGTAKRVVIDKENTTIVDGAGKSEDIKARVKQIRAQIBETSDDY 360  
Db 301 ISEBIGMELEKATLEDLQAKRVVINKDTTIIIDGVGEAAIQGRVAQIRQOIEBATSDDY 360  
Qy 361 DREKLOERLAKLVGVAVIHVGAATETEMKEKKORVEDALNATRAAEEGIVPGGTAFAV 420  
Db 361 DREKLOERKAVKLAGGVAVIKVGAATEVEMKEKKARVEDALHATRAAEEGVVAGGVALI 420  
Qy 421 RSIKVLDDIKPADDDDELAGNIIRSLSEEPRIQIAANAGYEGSVVEKVRPKDGFGENA 480  
Db 421 RVASKLADLRQONEDQNGIKVALRAMAPLRQIVLNCGEPSVVANTVKGDDNGYNA 480  
Qy 481 ASGEYEDLIKAGVIDPKKVTIRIALQNAASVASLLLTTECAIAEKPEPKKMPMPG--GGM 538  
Db 481 ATEEYGNMIDGILPTKVTIRSAQYAAVAGLMTTECMVTDL--PKNDAAADLGAAGGM 538  
Qy 539 GGMGMDGM 547  
Db 539 GGMGMDGM 547

RESULT 15  
US-08-470-260-6  
; Sequence 6, Application US/08470260  
; Patent No. 6077706  
; GENERAL INFORMATION:  
; APPLICANT: Covacci, Antonello  
; APPLICANT: Bugnoli, Massimo  
; APPLICANT: Telford, John  
; APPLICANT: Macchia, Giovanni  
; APPLICANT: Rappuoli, Rino  
; TITLE OF INVENTION: Helicobacter Pylori Proteins Useful  
; TITLE OF INVENTION: for Vaccines and Diagnostics  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/470,260  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/256,848  
; FILING DATE: 21-OCT-1994  
; ATTORNEY/AGENT INFORMATION:

; NAME: McClung, Barbara G.  
; REGISTRATION NUMBER: 33,113  
; REFERENCE/DOCKET NUMBER: 0316.001  
; TELEPHONE: (510) 601-2708  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 546 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-470-260-6  
  
Query Match 63.6%; Score 1723.5; DB 3; Length 546;  
Best Local Similarity 64.1%; Pred. No. 4.2e-142;  
Matches 350; Conservative 73; Mismatches 120; Indels 3; Gaps 2;  
  
Qy 3 SKEITFDKAREKLSRGVDKLANAVKTLGPKGRNVVIEKSFSPVITKDGVSVAKEI 62  
Db 2 AKETKFSDSARNLLFEGVRQLHDAVKVTMGPRGNVLQKSYGAPSIITKDGVSVAKEI 61  
  
Qy 63 EDKFNMGQMKVEKAPKTSIAGDGTATVLAQAIYREGVKLVAAAGRNPMIAIKRGI 122  
Db 62 SCPVANNMGALVKEVASKTADAAGDGTATVLAQAIYREGVKLVAAAGRNPMIAIKRGI 121  
  
Qy 123 AVAVTKELSDITKPRDQKEIAQVGTISANSDDTTGNIIAAMAKVKGKGVITVEEA 182  
Db 122 AAEAIINELKASKKVGKKEITQVATISANSDDTTGNIIAAMAKVKGKGVITVEEA 181  
  
Qy 183 LETTLDVVGKMFDRGYLSPYFVTNPEKVMVCELDNPEYILCNEKKITSMKMDLPIL 242  
Db 182 IEDELVDVVGKMFDRGYLSPYFVTNPEKVMVCELDNPEYILCNEKKITSMKMDLPIL 241  
  
Qy 243 VNRPLLIIEADVEGEALATLVNKLKRGALQVAVKAPGFGERRKAMLEDAIILTGE 302  
Db 242 EGKPLLIIEADVEGEALATLVNKLKRGALQVAVKAPGFGERRKAMLEDAIILTGE 301  
  
Qy 303 EDKFNMGQMKVEKAPKTSIAGDGTATVLAQAIYREGVKLVAAAGRNPMIAIKRGI 362  
Db 302 BELGLSLENAEVEFLGKAGRIVDKONTTIVDGKSHDDVDKRVQAIKTQIASITSDY 361  
  
Qy 363 EKLQERLAKLVGVAVIHVGAATETEMKEKKORVEDALNATRAAEEGIVPGGTAFAV 422  
Db 362 EKLQERLAKLVGVAVIKVGAATEVEMKEKKORVEDALNATRAAEEGIVPGGTAFAV 421  
  
Qy 423 IKVLDDIKPADDDDELAGNIIRSLSEEPRIQIAANAGYEGSVVEKVRPKDGFGENA 482  
Db 422 AQKVH--LNLHDDDEKGYEIIIMRAIKAPLAQIAINAGYDGVVNVVEKHEGHEG 479  
  
Qy 483 GEYEDLIKAGVIDPKKVTIRIALQNAASVASLLLTTECAIAEKPEPKKMPMPG--G 541  
Db 480 GKYVDFMEKEGIIDPLKVERIALQNAASVASLLLTTEATVHEIKEEKATPAMPD 539  
  
Qy 542 GGMGMDGM 547  
Db 540 GGMGMDGM 545

Search completed: April 12, 2005, 09:40:05  
Job time : 45 secs

**THIS PAGE BLANK (USPTO)**